Deeper Understanding: Addressing Methodological Constraints and Ethical Implications of Humanitarian Needs Assessments Using Natural Language Processing

Tino Kreutzer

A DISSERTATION SUBMITTED TO THE FACULTY OF GRADUATE STUDIES IN

PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

School of Health Policy & Management, York University

Graduate Program in Health

Supervisor: Prof. James Orbinski

13 April 2023

Abstract

Humanitarian needs assessments (HNAs) are used to prioritize and inform the provision of humanitarian assistance. HNA data are commonly collected by entering responses into computer-assisted personal interviewing (CAPI) tools. Even so, the proper transcription and translation of qualitative interview responses is hard to conduct rapidly and at scale during a humanitarian crisis, which has resulted in calls for deploying natural language processing (NLP), a type of artificial intelligence (AI) to augment CAPI tools. However, data about contemporary methodological HNA practices and problems, are lacking. At the same time, the risks to affected populations or to organizations providing humanitarian assistance from the processing of such data, including through the use of AI, are insufficiently understood. This manuscript-based study used a three-stage mixed-methods Design Science Research approach to inform the design and empirical evaluation of new features for KoboToolbox, a CAPI tool supported and used by humanitarian organizations, to systematically transcribe and translate answers to open-ended questions (OEQ) from HNA interviews. The first stage in the research was to conduct interviews with international key informants (KIs) to explore contemporary approaches to conducting HNA, as well as constraints related to different interview methods. The findings from interviews with 23 KIs from 13 countries revealed 47 themes. The second stage was a scoping review to map the range of ethical issues that have been raised in the academic literature related to processing data of people affected by humanitarian crises. The scoping review yielded 100 relevant studies after screening 8,387 papers, resulting in 22 themes. Third, new KoboToolbox software features were designed and implemented based on a user-centered design approach, and usability testing was conducted based on observations and feedback from test users. The two sets of themes from stages 1 and 2 were used to inform the iterative design and software implementation process,

which was validated through feedback from 14 test users from 10 countries. User testing showed strong support for the applicability and usability of new features and user testing results were subsequently used to identify and address several usability issues in the software. The tool resulting from this research addresses professional requirements and can now be readily deployed by humanitarian organizations globally to systematically transcribe, translate, and analyze answers to OEQs from HNA interviews.

Dedication

To my children, Theodore and Cléa

Acknowledgements

It has been a privilege to pursue my doctoral degree on a topic that I am so passionate about. The journey has been both challenging and rewarding, and I could not have made it this far without the support and encouragement of many people. It is my pleasure to express my heartfelt thanks to the many individuals who have helped me along the way.

First and foremost, I am most sincerely grateful to my supervisor, Professor James Orbinski, for his mentorship, conscientiousness, and unwavering support throughout the entire duration of my Ph.D. program. I am truly thankful for his patience, insight, and expertise, which have been invaluable in shaping my research and developing my academic skills. His thoughtful comments and detailed feedback have been crucial in shaping the direction of my work and improving the quality of my dissertation. I am truly indebted to him for his generosity, patience, and willingness to invest his time and energy in my academic growth.

Successful completion of this dissertation would not have been possible without my other committee members Professors Aijun An, Lora Appel, and Patrick Vinck. I am deeply grateful for their invaluable feedback, constructive criticism, and support throughout my doctoral journey. Their diverse perspectives, insights, and expertise helped me refine and strengthen my multidisciplinary research project. Their willingness to engage in rigorous discussions, provide insightful comments, and offer guidance and support throughout the many iterations of my dissertation project have made it possible to pursue research questions across such different academic areas.

I would like to thank York University, the Social Sciences and Humanities Research Council, the Vanier Canada Graduate Scholarship, the Dahdaleh Institute, and the Balsillie School of International Affairs for extending their financial support. I would also like to thank

Grand Challenges Canada for their generous funding, which has been instrumental in creating the innovations in KoboToolbox that were a central component of my research.

At York University, I would like to give particular thanks to the many people who supported me in various important ways behind the scenes over the years, including Prof. Farah Ahmad, Prof. Claudia Chaufan, Prof. Aryn Martin, Prof. Paul McDonald, Prof. Mazyar Fallah, Collette Murray, Theresa Dinh, and Syed Imran Ali.

I am grateful for the support and collaboration of my colleagues and friends, who have offered their invaluable feedback, encouragement, and support throughout the course of my research. Their diverse perspectives, insights, and friendship have made this journey a truly enriching and rewarding experience. I am indebted to my colleagues at Kobo and the Harvard Humanitarian Initiative, in particular Prof. Phuong Pham, Ignacio Rebollo, John Milner, and Vincenzo Bollettino, whose passion for survey work in humanitarian crises and pragmatic ideas have all contributed in important ways to the success of this dissertation.

My family's encouragement, love, and understanding have been a constant source of strength and motivation for me. I am grateful for their sacrifices, encouragement, and unconditional love, which have enabled me to pursue my academic dreams.

Last but not least, Maeve has been my steadfast rock throughout this journey. Through every challenge, she has been my partner, my sounding board, and my greatest advocate. Her unwavering faith in me and her willingness to listen to my ideas and concerns have been critical in shaping the direction of my research and my writing.

Table of Contents

List of Tables

List of Figures

List of Acronyms

Acronym	Definition
ACAPS	Formerly Assessment Capacities Project. Organization is now only referred to as ACAPS.
ACF	Action Contre la Faim
ASR	Automated speech recognition
CAPI	Computer-assisted personal interviewing
CATI	Computer-assisted telephone interviews
COVID-19	Coronavirus disease (2019)
DEEP	Data Entry and Exploration Platform
DRC	Danish Refugee Council
DTM	Data tracking matrix, an IOM program
GDPR	General Data Protection Regulation
HHI	Harvard Humanitarian Initiative
HNA	Humanitarian Needs Assessments
HRP	Humanitarian response plan
IASC	Inter-Agency Standing Committee
ICRC	International Committee of the Red Cross
IDP	Internally displaced person
IFRC	International Federation of Red Cross and Red Crescent Societies
iMMAP	Formerly Information Management and Mine Action Programs. Organization is now only referred to as iMMAP.
IMO	Information management officer
IOM	International Organization for Migration
IRC	International Rescue Committee
JIAF	Joint Intersectoral Analysis Framework
JIPS	Joint IDP Profiling Service
KI	Key informant
KII	Key informant interviews
MSF	Médecins Sans Frontières
MSNA	Multi-sectorial needs assessment
NGO	Non-governmental organization

Acronym	Definition
NLP	Natural language processing
NLU	Natural language understanding
OCHA	United Nations Office for the Coordination of Humanitarian Affairs
OEQ	Open-ended question
REACH	[REACH is not considered an acronym but the organization's name.]
SDR	Secondary data review
SRQR	Standards for Reporting Qualitative Research
STT	Speech-to-text
UCD	User-centered design
UNHCR	United Nations High Commissioner for Refugees; commonly branded as UN Refugee Agency
VAM	Vulnerability, Analysis and Mapping program by WFP
WFP	World Food Programme
WHO	World Health Organization
WHS	World Humanitarian Summit

Chapter 1 Introduction

> The resuscitation of a population is similar to the resuscitation of a severely injured patient, with needs assessment as the all important primary survey.
> (Redmond, 2005, p. 1320)

A young boy, washed up on a beach in Turkey; hospitals in Aleppo and Mariupol, obliterated from the air; an Ebola victim, turned away from the hospital to die in the streets of Monrovia: these haunting images illustrate the daunting challenges that humanitarian organizations face when trying to help people affected by a growing number of humanitarian crises globally. Organizations involved in providing humanitarian assistance work under strenuous circumstances and with limited funding to provide life-saving humanitarian assistance. However, resources for providing this assistance are far from sufficient. In 2022, donor governments provided US $ 27 billion to help 216 million people in 69 countries—a significantly smaller amount than the US $ 51.7 billion required to assist all people in need of humanitarian assistance for that year (OCHA, 2022, 2023). This considerable shortfall highlights the urgent need to better assess humanitarian needs and to do so at minimal cost.

The purpose of this dissertation is to investigate whether the design of new software features for KoboToolbox, a computer-assisted personal interviewing (CAPI) tool, might result in the ability to gather better data for humanitarian needs assessments (HNAs) to reflect the needs of affected people more accurately. KoboToolbox is an open-source CAPI software tool that was created in 2013 as a comprehensive platform for worldwide humanitarian data collection with support from the United Nations Office for the Coordination of Humanitarian

Affairs (OCHA) and in partnership with the UN High Commissioner for Refugees (UNHCR), and has been used by over 14,000 organizations globally to collect interview data (Kobo, 2022). This dissertation proposes that new software features should use natural language processing (NLP), a type of artificial intelligence (AI) to systematically transcribe and translate responses to open-ended questions, while also accounting for the ethical issues in developing and deploying such technology.

This manuscript-based dissertation uses a mixed-method Design Science Research approach rooted in pragmatist philosophy and uses key informant interviews, a scoping review of the literature, and the collection of observations of and feedback from test users about the tool's usability.

This introduction begins with an explanation of the background and rationale for the study, including definitions used in the dissertation. It then provides a literature review for three relevant research areas, the research questions, describes the ethical review of the study, and offers an overview of the contents of each of the dissertation chapters.

1.1. Background

Humanitarian assistance (sometimes also referred to as humanitarianism, humanitarian relief, humanitarian aid, or humanitarian action) as practiced by professionals and volunteers today, is typically seen as rooted in human history through expressions of solidarity and compassion (Barnett, 2011; Simms & Trim, 2011). These practices can be traced back to cultures in many parts of the world—such as, for example, East and Southeast Asia (Yeophantong, 2014), the Middle East (Moussa, 2014), or Latin America (Martínez, 2016). Some have described

humanitarian assistance as an “insurance of last resort” (Duffield, 2008, p. 151) to people without recourse to any regular state or private institutions due to war or disasters.

Beyond religious groups that have provided humanitarian assistance for centuries, multiple layers of national, intra-state, and international humanitarian NGOs have been established since the mid-nineteenth century as expressions of humanitarian solidarity (Barnett, 2013; Bass, 2008; Scholte, 2014). Professional humanitarian assistance was initially governed and provided almost exclusively by the International Committee of the Red Cross (ICRC). The ICRC was founded in 1863 to provide assistance to injured combatants from all sides of the conflict as well as, later, to civilian victims. The ICRC today is recognized in international law as the guardian of International Humanitarian Law (IHL), and is recognized by states by the granting of diplomatic status for its delegates.

Oxfam was founded in England in 1942 during World War II and was among the first formal non-governmental organizations (NGOs) involved in humanitarian assistance. Humanitarian NGOs do not have diplomatic status but claim legitimacy and protection for their actions in situations of armed conflict under IHL. The number of actors involved in providing humanitarian assistance increased significantly during and after World War II with the emergence of many NGOs and the formation of several United Nations Agencies dedicated to humanitarian relief and protection under IHL (Barnett, 2011). During the period of post-colonization conflicts, many NGOs were created with the intention to increase the amount of humanitarian assistance or to provide it in a way that was different from the accepted rules set by the UN or ICRC. While many NGOs remain limited in scope and regional in focus, some have grown to a substantial size: For example, a group of French physicians founded Médecins Sans

Frontières (MSF) in 1971, an organization that in 2021 had more than 63,000 staff globally and spent US $ 1.8 billion on its humanitarian operations (MSF, 2022).

As a result, the number of organizations and regimes responding to humanitarian crises has increased dramatically over the last 50 years to form an "institutional mosaic composed of multiple, partially overlapping institutional forms and regulatory configurations that are neither congruent, contiguous, nor coextensive with one another" (Brenner, 2004). While it remains central to this mosaic, the nation state is therefore only one of several actors at multiple scales, whereas humanitarian organizations can be at once implementers, decision makers, and public advocacy groups (Barnett, 2013). Although governments still provide the bulk of funding for humanitarian operations today, their influence is usually intentionally limited to ensure the independence of implementing organizations (United Nations, 2016).

Over the last two decades, humanitarian assistance has experienced four significant challenges. First, there has been a marked increase in the number of forcibly displaced people due to a combination of violent conflicts, food crises, disasters, and weather events. In 2022, there were about 103 million displaced people, a figure that was more than double that in 2010 (OCHA, 2022, p. 15). Disasters, which are responsible for the majority of short-term displacement (IDMC, 2022, p. 16), are being exacerbated by climate change which is increasing levels of vulnerability and risks to people and countries with the least capacity to respond to the growing number of climate-related disasters (IPCC, 2022).

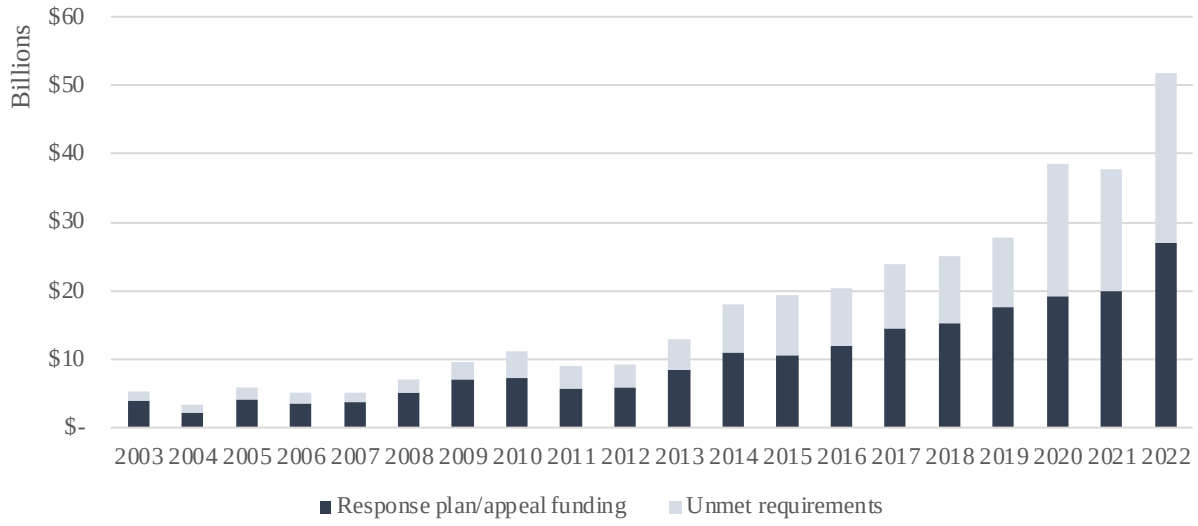
Second, the sheer scale of humanitarian needs has grown significantly. At the start of 2023, 339 million people are considered to be in need of humanitarian assistance (OCHA, 2022, p. 5), which represents a 44% increase from the start of 2021 (235 million) (OCHA, 2021a, p. 9). A considerable driver of this increase is what the United Nations has called the "largest global

food crisis in modern history" (OCHA, 2022, p. 22), with more than 222 million people facing acute food insecurity and being in need of urgent assistance at the end of 2022 (FAO & WFP, 2022, p. vii).

Third, there is a growing gap between humanitarian needs and available funding. Although absolute funding for United Nations coordinated appeals and response plans has increased significantly—especially since 2012—the proportion of unfunded needs has also grown in percentage terms, as shown for the last 20 years in Figure 1.

Figure 1

Humanitarian Response Plan/Appeal Funding and Unmet Requirements 2003-2022 (US $ billions) Based on Data Available from the United Nations Financial Tracking Service (OCHA, 2023)



Fourth, humanitarian organizations are also being increasingly targeted by violence, including due to attacks against humanitarian workers and relevant infrastructure. In 2021, 141 aid workers (Humanitarian Outcomes/Aid Worker Security Database, 2022) as well as 219 healthcare workers were killed in situations of armed conflict (United Nations, 2022). These

increased physical risks, which mark a decreasing respect for IHL, have led to a growing number of operations that are being led and monitored remotely. Such remote management of humanitarian assistance operations, including remotely-conducted HNA to understand population needs, has become even more widespread due to operational restrictions imposed by the COVID-19 pandemic.

Given the mosaic of different humanitarian actors and the confluence of significant challenges to the provision of humanitarian assistance described above, accurate HNA can play a crucial role in maximizing the effective and efficient delivery of humanitarian assistance. HNAs are an important first step toward providing humanitarian assistance because they provide critical evidence of population needs in order to guide decisions on the most appropriate interventions (Redmond, 2005). Mixed-methods approaches that combine qualitative and quantitative data collection methods are often considered optimal for informing relevant humanitarian assistance during in-depth assessments (OCHA, 2021b; UNHCR, 2017a).

Because of handheld CAPI applications developed in the twenty-first century, quantitative data can be collected more rapidly than ever before. Like paper-based methods, CAPI tools require interviewers or respondents themselves to convert often complex responses into categorical or numeric data that can then be objectively analyzed with relative ease and speed (Darcy, 2003). More nuanced insights often require the use of qualitative methods, but proper transcription and translation are hard to conduct rapidly and at scale during a crisis. Instead, most HNAs use predominantly structured surveys with closed-ended response options, that are quick to provide data, but that often fail to capture important qualitative nuances. At the same time, responses to open-ended questions (OEQ) often contain important contextual information not captured with quantitative methods. NLP methods can provide important new

opportunities to rapidly transcribe and translate voice responses for proper analysis to inform operational decision making in humanitarian crises. However, HNA professionals currently lack the tools and approaches to be able to take advantage of this quickly evolving technology (Emergency Data Science Workshop, 2018).

1.2. Rationale for this Study

The urgent need to identify better solutions for processing qualitative data in HNA was first discussed by several leading humanitarian organizations at the International Health Emergency Data Science Workshop at York University, 4-5 December 2018, based on a presentation given by the International Rescue Committee (IRC) that described the challenge of how to quickly and reliably analyze qualitative data at scale in an ongoing humanitarian crisis response (Emergency Data Science Workshop, 2018). A working group was then formed at the event consisting of humanitarian practitioners, design specialists, and researchers from ACAPS, Elrha, the Harvard Humanitarian Initiative (HHI), IRC, NetHope, OCHA, Pivotal, Purple Compass, the World Food Programme, and the Dahdaleh Institute for Global Health Research at York University. This dissertation author, a male researcher based in Canada who has also been serving as the Chief Operating Officer at Kobo, Inc, a non-profit humanitarian organization that develops KoboToolbox, was a co-organizer of the workshop and a member of the working group. Working group members further defined the problem by examining additional use cases and examples. Facilitated by a design specialist using participatory design methods, the group then identified potential innovations that would involve the systematic transcription, translation, and automated analysis of spoken interview responses for the purposes of HNA surveys using CAPI tools and feedback collection. Among these, the working group concluded that the use of NLP provided “potentially far-reaching new opportunities to capture qualitative data from voice

responses and analyze it for relevant content to better inform humanitarian assistance decisions," while also noting that ethical issues may arise from using such novel tools when processing data from vulnerable populations (Dahdaleh Institute for Global Health Research, 2019). Following these early ideas, the dissertation author obtained funding from Humanitarian Grand Challenges in 2020 to cover the professional design and software development work to address the above challenges, which would in turn be informed by the dissertation research.

This dissertation uses several terms that do not have universally accepted meanings, particularly in the humanitarian sector. These terms are defined in the following section for better understanding.

1.3. Definition of Terms

Humanitarian assistance is understood here to refer to coordinated actions that seek to "save lives and alleviate suffering of a crisis-affected population" (UN OCHA, 2003, p. 13). It also includes "protection," which "encompasses all activities aimed at obtaining full respect for the rights of the individual in accordance with the letter and the spirit of the relevant bodies of law" (IASC, 1999, p. 4).

Humanitarian crises are defined for the purposes of this study as a "series of events representing a critical threat to the health, safety, security or wellbeing of a community, usually over a wide area" (WHO, 2007, p. 7). Protracted humanitarian crises are defined as armed conflicts that persist over long periods of time (ICRC, 2016).

Humanitarian needs assessments are understood in this research to be the "set of activities necessary to understand a given situation, [which] entail the collection, updating and analysis of data pertaining to the population of concern (needs, capacities, resources, etc.), as

well as the state of infrastructure and general socio-economic conditions in a given location/area" (UNHCR, 2006, p. 4).

Structured interviews refer to data collection using standardized questionnaires in which primarily closed-ended questions are asked in a pre-determined order; whereas semi-structured interviews are approaches to data collection based on a list of primarily open-ended questions (OEQ) that are used to stimulate and guide discussions (ACAPS, 2014; Creswell, 2014).

For the purposes of this thesis, data processing is understood as: "Any operation or set of operations which is performed on data or on sets of data, whether or not by automated means, such as collecting, registering, storing, adapting or altering, cleaning, filing, retrieving, using, disseminating, transferring and retaining or destroying" (OCHA, 2019, p. 47).

Ethical issues are defined here as actions that may not conform to relevant moral standards, particularly those moral standards implicit in relevant humanitarian principles, because of the risks they present to affected populations or to organizations providing humanitarian assistance (Slim, 2015, pp. 47–145).

Artificial Intelligence (AI) is defined as scientific and technical attempts to build machines that act rationally, with the capacity to mimic rational cognitive functions to perceive, understand, predict, or manipulate (Russell & Norvig, 2009, pp. 1–30).

NLP is understood to be "the use of computational methods to analyze and process spoken or written statements in a language commonly used by humans" (Assal et al., 2011, p. 2).

Open-source software is "software with source code that anyone can inspect, modify, and enhance" (Opensource.com, n.d.) and that follows the requirements set by the Open Source Initiative, which include, among others, allowing anyone to research and inspect the software

code, enabling unlimited (including commercial) use, permitting anyone to modify the software, and distributing the software code without a licensing fee (Open Source Initiative, 2007).

The term usability describes “the extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use” (International Organization for Standardization, 2018).

User-centered design (UCD) is defined as a framework for flexibly and iteratively creating innovations that are supported by research to become user realities, as well as the institutional context in which innovations are intended to be used (Holden & Boustani, 2021). UCD approaches typically require 1) a thorough understanding of the eventual users of a new artifact (such as new software, goods, or services) in order to take their needs and environment into account, 2) empirical evaluations of user interactions, and 3) iterative cycles of designing and building new systems or services (Gould & Lewis, 1983).

1.4. Overview of the Literature

This section first provides an overview of the literature that relates to the methodological challenges of conducting HNAs to inform humanitarian assistance, particularly on the issues faced when collecting qualitative information. Second, it briefly outlines the ethical issues of processing data from people affected by humanitarian crises that are reported in the literature, while also describing how humanitarian organizations have responded to these issues. Third, this overview describes what has been published about the methods with which NLP could be used and integrated into a CAPI tool to improve HNA. Additional references to the literature are also included in each of the manuscript chapters to provide further details specific to each study.

1.4.1. Methodological Challenges Conducting HNA

HNAs typically provide critical evidence of population needs in order to guide decisions on the most appropriate interventions (Gerdin et al., 2014; Redmond, 2005), create a baseline for measuring impact (Banatvala, 2000; Spiegel et al., 2001), and help prioritize available funding (Cosgrave, 2009). The existing research shows that, depending on the phase of the humanitarian response, various methods are appropriate. Most HNAs are conducted through a combination of quantitative and qualitative methods (UNHCR, 2017a).

HNA data from structured interviews, mostly using quantitative methods, are commonly collected by entering responses into different kinds of handheld CAPI tools. For example, Building Markets and Orange Door Research (2018) has documented that KoboToolbox, is used by the majority of humanitarian organizations for HNA and other types of collecting primary data in humanitarian crises.

Guidance documents for humanitarian organizations show that qualitative methods in semi-structured or unstructured interviews are used in HNA to capture more nuanced insights, often by recording responses to OEQ in audio or text format for later analysis (for example, ACAPS, 2012). A review of such documents has shown that qualitative methods are usually preferred by humanitarian organizations during the first four weeks of a new or worsening humanitarian crisis, whereas a mixed-methods approach of using both qualitative and quantitative data collection methods is often considered optimal for informing relevant humanitarian assistance after the first month of the crisis, or in the case of protracted crises (OCHA, 2021b; UNHCR, 2017a).

As humanitarian staff from IRC at the *Emergency Data Science Workshop* (2018) indicated, there are numerous challenges that prevent organizations from using qualitative

methods for HNA at a larger scale. In practice, current HNA approaches often require interviewers to convert complex responses to OEQ in surveys or key informant interviews into simplified quantitative data or short notes. Figure 2 shows an example of a question in an HNA used by UNHCR that displays a long list of possible precoding response options. This is typically done with the use of structured questionnaires in CAPI tools by having interviewers code responses in the field, or by writing down short sentences that are then later coded during analysis (ACAPS, 2016). Without pre-coded options, the interviewer is required to rapidly take notes of what the respondent said, or record the response as an audio recording, which then needs to be replayed at a later stage for transcription, possible translation, and analysis. Regardless of the method of recording the response to OEQ, the diversity of languages and dialects common in many humanitarian crises can make collecting and analyzing qualitative data extremely difficult and costly (ACAPS, 2019; Al-Amer et al., 2016; Bowden & Fox-Rushby, 2003; Deutscher, 2010; Translators Without Borders, 2018a, 2018b, 2019).

Figure 2

Example Open-Ended Question from a Water and Sanitation Survey Used by UNHCR (2017b)

H4 (Op)/ Please tell me all the ways to prevent you or your household members from getting diarrhoea (*Do not prompt with responses, allow respondent to list and check those that are listed*)

- ☐ Boil or treat your water/drink clean water
- ☐ Wash hands with soap and water
- ☐ Cook food well
- ☐ Wash fruits and vegetables
- ☐ Cover food
- ☐ Cleaning cooking utensils
- ☐ Clean your home with bleach
- ☐ Use toilet/latrine facility to defecate
- ☐ Dispose of children's faeces in toilet/latrine
- ☐ Bury faeces
- ☐ Receive a vaccine
- ☐ Store water safely
- ☐ Breastfeeding babies
- ☐ Other: _____
- ☐ Don't know

Practitioner-level efforts to improve the quality and usability of HNA have included several recent initiatives. The 2016 Humanitarian Grand Bargain included several workstreams aimed at improving HNA (WHS, 2016), which resulted in a review of HNA analytical frameworks (Chataigner, 2017), a method of assessing the quality of HNAs (GPPI & INSPIRE Consortium, 2019; Okular Analytics, 2021), and the creation of the Joint Intersectoral Analysis Framework in 2021 to standardize HNA in protracted crises (OCHA, 2021b). Most of this work has focused primarily on improving the use of structured household-level data. To date, however, there has been very little focus on documenting the challenges of using qualitative methods in HNAs, from either an academic or a practitioner perspective.

Most available literature about HNA is published by humanitarian organizations themselves and generally focuses on how needs assessments should be conducted. There is little evidence of any study about the prevailing practices and challenges of applying these normative guidelines related to quantitative and qualitative methods, nor are there any studies found that empirically identify solutions to these challenges.

1.4.2. Ethical Issues Related to Processing Data from People Affected by a Humanitarian Crisis

Humanitarian organizations rely on processing increasingly large amounts of data to inform their operations (IASC, 2021), much of which is collected directly from affected populations through HNA or for other purposes, e.g., through refugee registrations, household surveys, or cash disbursements (e.g., Jacobsen, 2017; Thylin & Duarte, 2019; Vannini et al., 2020). As mentioned above, humanitarian staff are themselves increasingly victims of kidnappings and killings in conflict settings, which has led organizations to increasingly resort to remote methods of managing operations and collecting data from affected people (Donini &

Maxwell, 2013). This combination of factors has led to an exponential increase in the amount of personal data that is being collected distributed, stored, and analyzed in various locations around the world.

ICRC and Privacy International (2018) have shown that organizations turning to new or existing digital tools to collect, store, or analyze data more efficiently may knowingly or inadvertently introduce new ethical issues affecting people who are already vulnerable. As Sandvik et al. (2014) have argued, weighing the responsible use of new technologies in humanitarian crises comes with a number of ethical issues. Such issues may include, for example, confidential refugee registration data that may inadvertently be accessed by conflict parties (Jacobsen, 2015), distrust of new tools used in health emergency responses (Perakslis, 2018), or the loss of dignity due to intrusive use of drones in humanitarian assistance (van Wynsberghe & Comes, 2020). Ethical decisions are routinely made in humanitarian assistance about what data to collect, which tools to use, or how and with whom to share this information to avoid adverse consequences (Raymond et al., 2016; UN Global Pulse, 2016). In light of such challenges, organizations rarely choose to forego new tools altogether, such as when Oxfam decided to temporarily halt the use of biometrics in its programs (The Engine Room and Oxfam, 2018). Rather, many organizations are more likely to invest in new innovations without appropriately considering, weighing, or fully grasping the long-term ethical issues (Parker, 2019).

Because of these ethical challenges, several guidelines have been produced by humanitarian organizations with the goal of minimizing or eliminating ethical issues and the risks they present to affected populations or to organizations providing humanitarian assistance. Notable examples include *Data Responsibility in Humanitarian Action* by the Inter-Agency

Standing Committee (2021), the *Handbook on Data Protection in Humanitarian Action* by ICRC (2020), the *Data Responsibility Guidelines* by UN OCHA (2019), and the ethical design of new tools (Elhra & Humanitarian Health Ethics, 2021). Similarly, regulatory changes in many countries (such as the European Union's General Data Protection Regulation, GDPR) have moved many humanitarian organizations to change their approaches to data processing to address ethical issues related to data privacy (Bharania, 2017; Gazi, 2020). For example, Médecins Sans Frontières in France requires headquarters staff to attend mandatory GDPR training (MSF, 2019).

However, the speed of technological innovation as well as new circumstances (such as newly insecure environments or lack of access to populations during the COVID-19 pandemic) mean that guidance or regulation in a specific context may quickly become out of date and be inadequate to the new circumstances. Tegmark (2017) notes that AI systems that use machine learning and other methods for automating data processing are ushering in a completely new set of ethical issues that humanitarian organizations will have to confront. For example, some AI technologies were less accurate when transcribing speech to text from women (Garnerin et al., 2019; Tatman, 2017) or from African Americans (Koenecke et al., 2020). When relying on AI systems such issues can therefore mean that particular population groups could not be represented in survey data and thereby result in what Sarı et al. (2021) termed "fairness gaps."

Wang et al. (2021) argue that a distinct challenge for identifying relevant ethical issues is the lack of established ethical categories or theories used by studies discussing ethics in the humanitarian sector. First introduced by the International Federation of Red Cross and Red Crescent Societies (Pictet, 1979), the four humanitarian principles—humanity, impartiality, independence, and neutrality—are now widely and normatively used among many humanitarian

organizations (see, for example, MSF, n.d.), in international law (OCHA, 2012), as well as in ethical codes attempting to guide the actions of the humanitarian sector as a whole (see, for example, IFRC, 1994; Sphere Association, 2018). However, previous studies have shown the difficulty of applying these humanitarian principles in everyday practice (Hilhorst & Schmiemann, 2002), in guiding the use of information technology (Raymond & Card, 2015), or in mapping humanitarian organizations' ethical obligations (Broussard et al., 2019). In particular, Slim (2015) has argued that the broad *humanity* "principle" is better understood as an absolute moral value rather than an ethical principle (p. 62), whereas Orbinski (1994) has suggested that the humanitarian principles of independence and neutrality can be best understood as "mechanisms" (p. 7) for achieving the humanity principle. As a result, a growing number of studies use the ethical value categories of autonomy, beneficence, non-maleficence, and justice, which are widely used in the fields of bioethics and research ethics (Beauchamp & Childress, 2019), as better operational terms to reference ethical issues inherent to humanitarian practice (Cawthorne & Wynsberghe, 2019; Pham & Vinck, 2012).

To date, there is little evidence that there has been a comprehensive review of relevant ethical issues related to HNAs in the published literature. Previous systematic or scoping reviews focusing on humanitarian assistance only addressed more limited contexts, such as natural disasters (Tansey et al., 2018), and displaced populations (Makhoul et al., 2018), or did not include terms to capture more novel humanitarian activities such as responding to large-scale migration or public health emergencies (Pal et al., 2019). The gap in the literature is a catalyst for this current study.

1.4.3. Potential for and Methods of Integrating NLP in a CAPI tool to improve HNA

Galanis et al. (2021) point out that NLP has passed several important technological milestones over the last years, and Otter et al. (2021) conclude that this success is particularly due to breakthroughs in the areas of deep learning and neural networks. The rapid advances in these fields have led to an explosion of commercial NLP applications, particularly for automated speech recognition (ASR, and also known as speech to text, or STT) and machine translation (MT), as well as different types of content analysis, such as natural language understanding (NLU) or information extraction (Singh, 2018). Going from an audio recording to a translated written version of human speech typically requires first creating a written transcript that may then be translated into a target language (such as English)—though direct speech to translated text has also been achieved (Bansal et al., 2018; Weiss et al., 2017). High-quality transcription and translation software are now widely available through commercially available tools (e.g., from Google, IBM Watson, Microsoft Azure, Amazon Web Services). At the same time, open-source alternatives have been created to replicate or surpass the performances of these commercially available tools. Open Source examples are Transformer (Vaswani et al., 2017), DeepSpeech (Hannun et al., 2014), and OpenNMT (Klein et al., 2017).

But despite this progress and a clear need to improve multilingual communication, there is little evidence based on the available literature about the adoption of these tools in the humanitarian sector to date. One important factor may be that the varying accuracy of ASR and MT requires human correction—especially given the wide range of speakers, accents, recording quality, as well as the sheer novelty of the technology (Dew et al., 2018). A second challenge, particularly for HNAs, is that in some countries currently affected by humanitarian crises, commercial tools for ASR and MT do not support the local languages, such as Amharic, Hausa,

Rohingya, Swahili, or Fulfulde (Abbott & Martinus, 2018; Gu et al., 2018). As a potential alternative, two open-source software tools for ASR exist that can be used to create high quality ASR models without the use of commercial providers: Coqui (2022) and Mozilla's (2022b) DeepSpeech. These tools can be trained on training data for any language, including using freely available corpuses of language libraries, such as Mozilla's (2022a) CommonVoice and Tatoeba (2022). However, given that such public repositories tend to favor languages spoken in high and middle income countries, Translators Without Borders (Ansari & Petras, 2018) has begun investing in ASR and MT models specifically for languages spoken in countries that have suffered some of the most protracted humanitarian crises.

Apart from the technical possibilities, Davis (1989) shows that software ultimately needs to be designed in a way that is both useful and easy to use. Without proper focus on usability, many software innovations fail to deliver the expected impact despite considerable financial resources—particularly for publicly funded projects (Anthopoulos et al., 2016; Savoldelli et al., 2014). Best practices in web and software design therefore follow the user-centered design (UCD) approach. UCD is a widely-used framework for flexibly and iteratively creating innovations that are supported by research into user realities as well as the institutional context in which innovations are intended to be used (Holden & Boustani, 2021). UCD approaches typically require 1) a thorough understanding of potential users in order to take their needs and environment into account, 2) empirical evaluations of user interactions to measure usability, and 3) iterative cycles of designing and building new systems or services (Gould & Lewis, 1983). Failures to apply UCD during the design and development process can mean that new tools are rejected by organizations or professionals, or have the potential to cause harm (Cornet et al., 2019). These challenges are unique and acute in critical settings such as humanitarian assistance,

or in providing healthcare to vulnerable populations by often overstretched humanitarian staff (Ben-Zeev et al., 2015; Kushniruk & Nøhr, 2016). Likewise, whereas UCD is often embraced in theory, evaluating new tools empirically, based on user testing, has many challenges that can dissuade humanitarian innovators from embracing this practice more widely (Cornet et al., 2020).

At the time of the research, there is little evidence of any extant studies describing the systematic design for applying NLP technology to improve HNAs—while also avoiding usability issues and ethical issues in the process.

In summary, this literature review identified three distinct gaps in the literature:

1. **Prevailing practices and methodological challenges in HNA**: Despite the critical potential of HNA for informing life-saving humanitarian assistance and for prioritizing limited funding, very little research exists about prevailing practices, operational constraints, or methodological challenges in conducting HNA.
2. **Ethical issues in the design of HNA data processing tools:** Designing new data processing tools for HNA can lead to many ethical issues, but there is no systematic review of what these issues are or how these can be categorized.
3. **Ethically appropriate design of new CAPI features using NLP technologies:** NLP has had many recent success stories in research and commercial applications, but to date there are no studies describing the systematic design for applying this technology to improve HNA while avoiding usability issues that would limit its implementation or minimizing ethical issues that would present

risks to affected populations or to organizations providing humanitarian assistance.

1.5. Research Questions

There have been insufficient studies of these three gaps. For each of those three areas, research questions were developed.

Regarding the first gap of prevailing practices and methodological challenges in HNA, three specific research questions were identified:

1. What are the current approaches to conducting HNAs in a sample of professional staff members of humanitarian organizations with prior experience managing HNA data collection?
2. Among these respondents, what constraints related to different interview methods, particularly qualitative ones, are identified?
3. Which solutions for improving the use of qualitative interview methods in HNAs do respondents propose?

Regarding the second area related to ethical issues in the design of HNA data processing tools**,** three specific research questions were identified:

1. Which ethical issues have been raised in the peer-reviewed literature related to processing data from people affected by humanitarian crises in order to inform humanitarian assistance?
2. To what extent do real-world examples of ethical issues reflect the concerns presented in the literature?
3. Which technologies were the focus of concern over these ethical issues?

Regarding the third area of study related to the ethically appropriate design of new CAPI features using NLP technologies, three specific research questions were identified:

1. Which specific constraints identified by professionals with experience managing HNA data collection could be addressed through new CAPI features?
2. Which ethical issues identified in the peer-reviewed literature related to processing data from people affected by humanitarian crises should be considered when designing new CAPI features for processing qualitative data?
3. How should new features for the KoboToolbox CAPI software to systematically transcribe, translate, and analyze answers to OEQ from HNA interviews be designed, usability tested, and implemented?

1.6. Ethical Review of the Study

Ethics approval for the study was granted by the Office of Research Ethics at York University, Canada (certificate # STU 2020-092). Informed consent was obtained from all participants. Great care was taken to protect the anonymity of respondents. Quotations from key informants in chapters 4 and 6 were therefore not attributed to their organization or country of residency.

1.7. Dissertation Overview

This dissertation includes seven chapters, which are detailed below.

Chapter 1 (this introduction) provides the background, the rationale for this study, important definitions, an overview of the literature, the research questions, describes the ethical review of the study, and offers an overview of the structure and contents of the dissertation.

Chapter 2 describes the theoretical framework and provides an overview of the methodologies applied. It describes the use of the Design Science Research approach and how this dissertation is grounded in a pragmatist philosophy. It also outlines the values that have guided this research from the fields of humanitarian assistance, bioethics, and design science, respectively.

The findings of this dissertation research are laid out in four interconnected studies in chapters 3, 4, 5, and 6, using a format that is intended for publication in peer-reviewed academic journals. Of these, one manuscript (chapter 3) has already been published during the course of the dissertation research. As applicable, individual chapter manuscripts include a detailed review of the literature, a full description of the methods used, findings, discussions, as well as conclusions with indications on proposed next steps for research and practice. Whereas these chapters can be considered on their own, together and in sequence they provide a new perspective into how the areas of humanitarian assistance, ethics, and software development can and should be considered together under a Design Science Research approach. The following briefly describes each chapter.

Chapter 3 is a chapter manuscript entitled *Improving Humanitarian Needs Assessments through Natural Language Processing*, which was published in 2020 in *IBM Research and Development* (Kreutzer et al., 2020). The dissertation researcher conceived, designed, researched and wrote in its entirety the first draft of the manuscript chapter, with subsequent contributions from dissertation committee members, researchers with expertise in NLP and survey methods, and experts in humanitarian theory and practice. The co-authors are, in the order listed in the published version: Patrick Vinck, Phuong Pham, Aijun An, Lora Appel, Eric Deluca, Grace Tang, Muath Alzghool, Kusum Hachhethu, Bobi Morris, John Crowley, and James Orbinski.

Some of their prior research provided background for the design idea that inspired the current study. This chapter also describes the anticipated practical and ethical challenges of building on the diffusion of digital data collection platforms and introducing this new technology to the humanitarian assistance sector. It concludes by providing an overview of the ethical values or principles that should be used to anticipate and mitigate ethical issues.

Chapter 4 is a chapter manuscript titled *Overcoming Blind Spots: Constraints and Solutions Related to Qualitative Interview Methods in Humanitarian Needs Assessments*. Based on in-depth interviews conducted with a sample of international key informants (KIs) with prior experience managing HNA data collection, this study explores current approaches to conducting HNA, constraints related to different interview methods—particularly qualitative ones, and solutions for improving the use of qualitative interview methods in HNAs. 23 KIs were selected from 13 countries, representing 17 humanitarian organizations. Based on a grounded theory approach of analyzing interview transcript, the manuscript describes prevailing approaches to data collection in HNAs, constraints related to quantitative and especially qualitative interview methods, challenges specific to handling open-ended questions, and specific recommendations for improving the role of qualitative interview methods. Respondents pointed to a complex operational environment that favors quantitative methods in HNA, even though qualitative data are widely seen as crucial for understanding population needs. Among several concrete suggestions, KIs proposed acknowledging the limitations of quantitative data in HNA, investing in more qualitative skills through training or hiring, piloting mixed-methods hybrid surveys, as well as employing NLP for transcribing, translating, and analyzing interview responses to OEQ.

Chapter 5 is a chapter manuscript entitled *Ethical Implications Related to Processing of Personal Data in Humanitarian Crises: A Scoping Review*. This chapter manuscript mapped the

range of ethical issues that have been raised in the academic literature regarding the processing of data from people affected by humanitarian crises, based on a systematic search of databases to identify peer-reviewed studies published since 2010. The study used the scoping review method established by Arksey and O'Malley (2005) as further refined by Levac et al. (2010), followed the framework maintained by the Joanna Briggs Institute (Peters et al., 2017), and applied a study protocol using the PRISMA-ScR reporting guidelines for scoping reviews (Tricco et al., 2018). The scoping review yielded 100 relevant studies after screening 8,387 papers. A detailed review of the 100 studies led to the identification of 22 ethical issues which were grouped into four ethical value categories of autonomy, beneficence, non-maleficence, and justice, which were identified as the best method of categorizing such issues. The chapter manuscript demonstrated the need for specific and robust guidance to better navigate emerging ethical concerns. It also showed that further empirical research related to ethical issues is needed in conflict settings, specific to the use of CAPI tools, and linked to early adoptions of artificial intelligence in humanitarian assistance.

Chapter 6 is a manuscript titled *Systematic Design and Evaluation of New Humanitarian Needs Assessments Tools for Collecting Qualitative Data Using Natural Language Processing*. This chapter manuscript reports on the design and empirical evaluation of new KoboToolbox features to systematically transcribe, translate, and analyze answers to OEQ from HNA interviews. It first does so by extracting and analyzing a subset of 25 relevant themes identified in Chapter 4 from interviews with KIs, as well as by identifying a list of seven ethical issues from the scoping review presented in Chapter 5. Both sets of data are used to inform the initial design proposal of the new features. The study then describes the iterative user-centered design approach for creating the new software features and reports on the results from usability testing

based on observations and feedback from 14 test users. Results from user testing showed strong support for the usefulness and usability of new features. User testing results were subsequently used to identify several usability issues that were then addressed in the software. This study demonstrates that the adoption of a multi-disciplinary user-centered design approach was crucial for creating a new set of tools for scaling up the use of qualitative data in humanitarian needs assessments while accommodating both the ethical realities and the operational constraints experienced in humanitarian assistance.

Chapter 7 provides an overall summation and conclusions to this dissertation. It summarizes the dissertation's strengths and limitations, highlights the contributions to new knowledge and significance of the dissertation, suggests implications for further research, and offers specific recommendations for policy and practice.

1.8. References

Abbott, J. Z., & Martinus, L. (2018). Towards Neural Machine Translation for African Languages. *ArXiv:1811.05467 [Cs, Stat]*. https://arxiv.org/abs/1811.05467v1

ACAPS. (2012). *Qualitative and quantitative research techniques for humanitarian needs assessment*. ACAPS. https://www.acaps.org/sites/acaps/files/resources/files/qualitative_and_quantitative_research_techniques_for_humanitarian_needs_assessment-an_introductory_brief_may_2012.pdf

ACAPS. (2014). Humanitarian Needs Assessment: The Good Enough Guide. In *Geneva: ACAPS*. ACAPS. https://doi.org/10.3362/9781780449357

ACAPS. (2016). *Meeting information needs? A review of ten years of multisector coordinated needs assessment reports*. Geneva: ACAPS. https://www.acaps.org/sites/acaps/files/resources/files/acaps_report-a_review_of_ten_years_of_multisector_coordinated_needs_assessment_reports_january_2016.pdf

ACAPS. (2019). *Lessons Learned: Needs Assessments in Cox's Bazar*. ACAPS. https://www.acaps.org/sites/acaps/files/products/files/20190404_acaps_npm_coxs_bazar_analysis_hub_needs_assessments_lessons_learned.pdf

Al-Amer, R., Ramjan, L., Glew, P., Darwish, M., & Salamonson, Y. (2016). Language translation challenges with Arabic speakers participating in qualitative research studies. *International Journal of Nursing Studies*, *54*, 150–157. https://doi.org/10.1016/j.ijnurstu.2015.04.010

Ansari, A., & Petras, R. (2018). *Gamayun: The Language Equality Initiative*. Translators Without Borders. https://translatorswithoutborders.org/wp-

content/uploads/2018/03/Gamayun-Language-Equality-Initiative-March-2018.pdf

Anthopoulos, L., Reddick, C. G., Giannakidou, I., & Mavridis, N. (2016). Why e-government projects fail? An analysis of the Healthcare.gov website. *Government Information Quarterly*, *33*(1), 161–173. https://doi.org/10.1016/j.giq.2015.07.003

Arksey, H., & O'Malley, L. (2005). Scoping studies: Towards a methodological framework. *International Journal of Social Research Methodology*, *8*(1), 19–32. https://doi.org/10.1080/1364557032000119616

Assal, H., Seng, J., Kurfess, F., Schwarz, E., & Pohl, K. (2011). Semantically-enhanced information extraction. *2011 Aerospace Conference*, 1–14. https://doi.org/10.1109/AERO.2011.5747547

Banatvala, N. (2000). Conflict and health: Public health and humanitarian interventions: developing the evidence base. *BMJ*, *321*(7253), 101–105. https://doi.org/10.1136/bmj.321.7253.101

Bansal, S., Kamper, H., Livescu, K., Lopez, A., & Goldwater, S. (2018). Low-resource speech-to-text translation. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, *2018-Septe*, 1298–1302. https://doi.org/10.21437/Interspeech.2018-1326

Barnett, M. N. (2011). *Empire of humanity: A history of humanitarianism*. Cornell University Press.

Barnett, M. N. (2013). Humanitarian Governance. *Annual Review of Political Science*, *16*(1), 379–398. https://doi.org/10.1146/annurev-polisci-012512-083711

Bass, G. J. (2008). *Freedom's battle: the origins of humanitarian intervention*. Alfred A. Knopf.

Beauchamp, T. L., & Childress, J. F. (2019). *Principles of Biomedical Ethics* (7th editio). Oxford

University Press.

Ben-Zeev, D., Schueller, S. M., Begale, M., Duffecy, J., Kane, J. M., & Mohr, D. C. (2015). Strategies for mHealth Research: Lessons from 3 Mobile Intervention Studies. *Administration and Policy in Mental Health and Mental Health Services Research*, *42*(2), 157–167. https://doi.org/10.1007/s10488-014-0556-2

Bharania, R. (2017). *Exploring the Intersection of Technology and Emergencies The GDPR is a unique opportunity to get humanitarian data protection right* . The Digital Responder. https://thedigitalresponder.wordpress.com/2017/12/30/the-gdpr-is-a-unique-opportunity-to-get-humanitarian-data-protection-right/

Bowden, A., & Fox-Rushby, J. A. (2003). A systematic and critical review of the process of translation and adaptation of generic health-related quality of life measures in Africa, Asia, Eastern Europe, the Middle East, South America. *Social Science & Medicine*, *57*, 1289–1306. https://doi.org/10.1016/S0277-9536(02)00503-8

Brenner, N. (2004). *New state spaces : urban governance and the rescaling of statehood*. Oxford University Press.

Broussard, G., Rubenstein, L. S., Robinson, C., Maziak, W., Gilbert, S. Z., & DeCamp, M. (2019). Challenges to ethical obligations and humanitarian principles in conflict settings: a systematic review. *Journal of International Humanitarian Action*, *4*(1). https://doi.org/10.1186/s41018-019-0063-x

Building Markets, & Orange Door Research. (2018). *What is the Point if Nothing Changes: Current Practices and Future Opportunities to Improve Remote Monitoring and Evaluation in Syria*. New York, NY: Building Markets.

Cawthorne, D., & Wynsberghe, A. R. Van. (2019). From HealthDrone to FrugalDrone: Value-

Sensitive Design of a Blood Sample Transportation Drone. *International Symposium on Technology and Society, Proceedings*, *2019-Novem*. https://doi.org/10.1109/ISTAS48451.2019.8938019

Chataigner, P. (2017). *Analytical Framework Review report- A framework analysis of analytical frameworks*. OCHA. https://assessments.hpc.tools/km/analytical-framework-review-report

Coqui. (2022). *Coqui*. https://coqui.ai/

Cornet, V. P., Daley, C., Bolchini, D., Toscos, T., Mirro, M. J., & Holden, R. J. (2019). Patient-centered Design Grounded in User and Clinical Realities: Towards Valid Digital Health. *Proceedings of the International Symposium on Human Factors and Ergonomics in Health Care*, *8*(1), 100–104. https://doi.org/10.1177/2327857919081023

Cornet, V. P., Toscos, T., Bolchini, D., Ghahari, R. R., Ahmed, R., Daley, C., Mirro, M. J., & Holden, R. J. (2020). Untold stories in user-centered design of mobile health: Practical challenges and strategies learned from the design and evaluation of an app for older adults with heart failure. *JMIR MHealth and UHealth*, *8*(7), 1–14. https://doi.org/10.2196/17703

Cosgrave, J. (2009). Humanitarian funding and needs assessment. In *Humanitarian Response Index 2008: Donor Accountability in Humanitarian Action* (pp. 75–88). https://doi.org/10.1057/9780230584617

Creswell, J. W. (2014). *Research Design* (4th ed.). Sage.

Dahdaleh Institute for Global Health Research. (2019, July 15). *Natural Language Processing Project Takes Shape*. https://www.yorku.ca/dighr/natural-language-processing-project-takes-shape/

Darcy, J. (2003). *Measuring humanitarian need: A critical review of needs assessment practice and its influence on resource allocation*. https://www.odi.org/sites/odi.org.uk/files/odi-

assets/publications-opinion-files/4899.pdf

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly: Management Information Systems*, *13*(3), 319–339. https://doi.org/10.2307/249008

Deutscher, G. (2010). *Through the Language Glass: Why the World Looks Different in Other Languages*. Henry Hold and Company.

Dew, K. N., Turner, A. M., Choi, Y. K., Bosold, A., & Kirchhoff, K. (2018). Development of machine translation technology for assisting health communication: A systematic review. *Journal of Biomedical Informatics*, *85*(July), 56–67. https://doi.org/10.1016/j.jbi.2018.07.018

Donini, A., & Maxwell, D. (2013). From face-to-face to face-to-screen: Remote management, effectiveness and accountability of humanitarian action in insecure environments. *International Review of the Red Cross*, *95*(890), 383–413. https://doi.org/10.1017/S1816383114000265

Duffield, M. (2008). Global Civil War: The Non-Insured, International Containment and Post-Interventionary Society. *Journal of Refugee Studies*, *21*(2), 145–165. https://doi.org/10.1093/jrs/fem049

Elhra, & Humanitarian Health Ethics. (2021). *Ethics for humanitarian innovation: Toolkit*. https://higuide.elrha.org/wp-content/uploads/2021/09/Elrha-Ethics_for_Humanitarian_Innovation-Toolkit-210913.pdf

Emergency Data Science Workshop. (2018). *Natural Language Processing for Humanitarian Survey Work: Challenge Presented by International Rescue Committee*. https://emergencydatascience.org/challenge_nlp_surveys/

FAO, & WFP. (2022). *Hunger Hotspots: FAO - WFP early warnings October 2022 to January 2023 Outlook*. https://reliefweb.int/report/world/hunger-hotspots-fao-wfp-early-warnings-acute-food-insecurity-october-2022-january-2023-outlook

Galanis, N. I., Vafiadis, P., Mirzaev, K. G., & Papakostas, G. A. (2021). Machine Learning Meets Natural Language Processing - The Story so Far. *IFIP Advances in Information and Communication Technology*, *627*, 673–686. https://doi.org/10.1007/978-3-030-79150-6_53

Garnerin, M., Rossato, S., & Besacier, L. (2019). Gender representation in French broadcast corpora and its impact on ASR performance. *AI4TV 2019 - Proceedings of the 1st International Workshop on AI for Smart TV Content Production, Access and Delivery, Co-Located with MM 2019*, 3–9. https://doi.org/10.1145/3347449.3357480

Gazi, T. (2020). Data to the rescue: how humanitarian aid NGOs should collect information based on the GDPR. *Journal of International Humanitarian Action*, *5*(1). https://doi.org/10.1186/s41018-020-00078-0

Gerdin, M., Chataigner, P., Tax, L., Kubai, A., & von Schreeb, J. (2014). Does need matter? Needs assessments and decision-making among major humanitarian health agencies. *Disasters*, *38*(3), 451–464. https://doi.org/10.1111/disa.12065

Gould, J. D., & Lewis, C. (1983). Designing for usability-key principles and what designers think. *Conference on Human Factors in Computing Systems - Proceedings*, *28*(3), 50–53. https://doi.org/10.1145/800045.801579

GPPI, & INSPIRE Consortium. (2019). *Methodology to Assess Coordinated Multi-Sector Needs Assessments*. Global Public Policy Institute. https://www.gppi.net/project/assessing-needs-assessments

Gu, J., Wang, Y., Chen, Y., Cho, K., & Li, V. O. K. (2018). *Meta-Learning for Low-Resource*

Neural Machine Translation. http://arxiv.org/abs/1808.08437

Hannun, A., Case, C., Casper, J., Catanzaro, B., Diamos, G., Elsen, E., Prenger, R., Satheesh, S., Sengupta, S., Coates, A., & Ng, A. Y. (2014). *Deep Speech: Scaling up end-to-end speech recognition*. http://arxiv.org/abs/1412.5567

Hilhorst, D., & Schmiemann, N. (2002). Development in Practice Humanitarian principles and organisational culture: Everyday practice in Médecins Sans Frontières-Holland. *Development in Practice*, *12*(3–4), 490–500. https://doi.org/10.1080/0961450220149834

Holden, R. J., & Boustani, M. A. (2021). Design and Usability Methods. In R. J. Holden & R. S. Valdez (Eds.), *The Patient Factor* (1st ed.). CRC Press.

Humanitarian Outcomes/Aid Worker Security Database. (2022). *Aid Worker Security Report: Figures at a glance*. Humanitarian Outcomes. https://www.humanitarianoutcomes.org/sites/default/files/publications/awsd_figures_2022.pdf

IASC. (1999). *Protection of Internally Displaced Persons*. https://interagencystandingcommittee.org/system/files/legacy_files/protection_of_internally_displaced_persons_inter_agency_standing_committee_policy_paper_0.pdf

IASC. (2021). *Operational Guidance: Data Responsibility in Humanitarian Action*. Inter-Agency Standing Committee. https://interagencystandingcommittee.org/operational-response/iasc-operational-guidance-data-responsibility-humanitarian-action

ICRC. (2016). *Protracted conflict and humanitarian action | ICRC*. International Committee of the Red Cross. https://www.icrc.org/en/document/protracted-conflict-and-humanitarian-action

IDMC. (2022). *2022 Global Report on Internal Displacement*. Internal Displacement Monitoring

Centre. https://www.internal-displacement.org/sites/default/files/publications/documents/IDMC_GRID_2022_LR.pdf

IFRC. (1994). *The Code of Conduct for the International Red Cross and Red Crescent Movement and Non Governmental Organisations (NGOs) in Disaster Relief*. http://www.ifrc.org/Global/Publications/disasters/code-of-conduct/code-english.pdf

International Committee of the Red Cross. (2020). *Handbook on Data Protection in Humanitarian Action (2nd Edition)*. https://www.icrc.org/en/data-protection-humanitarian-action-handbook

International Committee of the Red Cross and Privacy International. (2018). *The Humanitarian Metadata Problem: "Doing No Harm" in the Digital Era*. https://privacyinternational.org/sites/default/files/2018-12/The Humanitarian Metadata Problem - Doing No Harm in the Digital Era.pdf

International Organization for Standardization. (2018, March). *Ergonomics of human-system interaction — Part 11: Usability: Definitions and concepts*. ISO 9241-11:2018(En). https://kebs.isolutions.iso.org/obp/ui#iso:std:iso:9241:-11:ed-2:v1:en

IPCC. (2022). Summary for Policymakers. In H. O. Pörtner, D. C. Roberts, M. Tignor, E. S. Poloczanska, K. Mintenbeck, A. Alegría, M. Craig, S. Langsdorf, S. Löschke, V. Möller, A. Okem, & B. Rama (Eds.), *Climate Change 2022: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (p. In Press). Cambridge University Press. https://www.ipcc.ch/report/ar6/wg2/downloads/report/IPCC_AR6_WGII_SummaryForPolicymakers.pdf

Jacobsen, K. L. (2015). Experimentation in humanitarian locations: UNHCR and biometric

registration of Afghan refugees. *Security Dialogue*, *46*(2), 144–164. https://doi.org/10.1177/0967010614552545

Jacobsen, K. L. (2017). On Humanitarian Refugee Biometrics and New Forms of Intervention. *Journal of Intervention and Statebuilding*, *11*(4), 529–551. https://doi.org/10.1080/17502977.2017.1347856

Klein, G., Kim, Y., Deng, Y., Senellart, J., & Rush, A. M. (2017). *OpenNMT: Open-Source Toolkit for Neural Machine Translation*. http://arxiv.org/abs/1701.02810

Kobo. (2022). *About the Kobo Organization*. https://www.kobotoolbox.org/about-us/the-organization/

Koenecke, A., Nam, A., Lake, E., Nudell, J., Quartey, M., Mengesha, Z., Toups, C., Rickford, J. R., Jurafsky, D., & Goel, S. (2020). Racial disparities in automated speech recognition. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(14), 7684–7689. https://doi.org/10.1073/pnas.1915768117

Kreutzer, T., Vinck, P., Pham, P., An, A., Appel, L., Deluca, E., Tang, G., Alzghool, M., Hachhethu, K., Morris, B., Crowley, J., & Orbinski, J. (2020). Improving Humanitarian Needs Assessments through Natural Language Processing. *IBM Journal of Research and Development*, *64*(1). https://doi.org/10.1147/JRD.2019.2947014

Kushniruk, A., & Nøhr, C. (2016). Participatory design, user involvement and health IT evaluation. *Evidence-Based Health Informatics: Promoting Safety and Efficiency through Scientific Methods and Ethical Policy*, *222*, 139–151. https://doi.org/10.3233/978-1-61499-635-4-139

Levac, D., Colquhoun, H., & O'Brien, K. K. (2010). Scoping studies: Advancing the methodology. *Implementation Science*, *5*(1). https://doi.org/10.1186/1748-5908-5-69

Makhoul, J., Chehab, R. F., Shaito, Z., & Sibai, A. M. (2018). A scoping review of reporting "Ethical Research Practices" in research conducted among refugees and war-affected populations in the Arab world. *BMC Medical Ethics*, *19*(1), 1–9. https://doi.org/10.1186/s12910-018-0277-2

Martínez, L. L. (2016). Pueblos indígenas y acción humanitaria en América Latina y el Caribe. In I. Mosel, C. Bennett, & H. Krebs (Eds.), *Aproximaciones a la historia del humanitarismo en América Latina y el Caribe*. ODI/HPG. https://cdn.odi.org/media/documents/Aproximaciones_a_la_historia_del_humanitarismo_en_América_Latina_y_el_Caribe.pdf

Moussa, J. (2014). *Ancient origins, modern actors: Defining Arabic meanings of humanitarianism*. ODI/HPG. https://cdn.odi.org/media/documents/9290.pdf

Mozilla. (2022a). *Common Voice*.

Mozilla. (2022b). *mozilla/DeepSpeech*. https://github.com/mozilla/DeepSpeech

MSF. (n.d.). *Independent, impartial, neutral*. Retrieved January 4, 2021, from https://www.doctorswithoutborders.org/who-we-are/principles/independence

MSF. (2019). *The Data Protection Course (OCP)*. MSF Tembo. https://tembo.msf.org/course/info.php?id=265

MSF. (2022). *Financial Report 2021*. https://www.msf.org/sites/default/files/2022-05/MSF_Financial_Report_2021_FINAL provisional.pdf

OCHA. (2012). *Humanitarian Principles*. UN Office for the Coordination of Humanitarian Affairs. https://www.unocha.org/sites/dms/Documents/OOM-humanitarianprinciples_eng_June12.pdf

OCHA. (2019). *Data responsibility guidelines*. UN Office for the Coordination of Humanitarian

Affairs. https://centre.humdata.org/wp-content/uploads/2019/03/OCHA-DR-Guidelines-working-draft-032019.pdf

OCHA. (2021a). *Global Humanitarian Overview 2022*. Office for the Coordination of Humanitarian Affairs. https://doi.org/10.18356/9789210012423

OCHA. (2021b). *Joint intersectoral analysis framework 1.1*. UNOCHA. https://doi.org/10.18356/9789214030751c020

OCHA. (2022). *Global Humanitarian Overview 2023*. Office for the Coordination of Humanitarian Affairs. https://reliefweb.int/report/world/global-humanitarian-overview-2023-enaresfr

OCHA. (2023). *Financial Tracking Service*. https://fts.unocha.org

Okular Analytics. (2021). *Technical Review of 2020 MSNAs in Protracted Crises*. IMPACT Initiatives. https://www.reachresourcecentre.info/wp-content/uploads/2021/07/Okular_Report_MSNAs2020ExternalReview_July2021.pdf

Open Source Initiative. (2007, March 22). *The Open Source Definition*.

Opensource.com. (n.d.). *What is open source?* Retrieved May 1, 2022, from https://opensource.com/resources/what-open-source

Orbinski, J. (1994). Health Impact of Rights Violations in Conf1ict Zones and the Humanitarian Imperative: Rwanda-A Case Study. *Refuge: Canada's Journal on Refugees*, *14*(5), 5–8. https://doi.org/10.25071/1920-7336.21827

Otter, D. W., Medina, J. R., & Kalita, J. K. (2021). A Survey of the Usages of Deep Learning for Natural Language Processing. *IEEE Transactions on Neural Networks and Learning Systems*, *32*(2), 604–624. https://doi.org/10.1109/TNNLS.2020.2979670

Pal, N. E., Eckenwiler, L., Hyppolite, S.-R., Pringle, J., Chung, R., & Hunt, M. (2019). Ethical

considerations for closing humanitarian projects: a scoping review. *Journal of International Humanitarian Action*, *4*(1). https://doi.org/10.1186/s41018-019-0064-9

Parker, B. (2019, February 5). *New UN deal with data mining firm Palantir raises protection concerns*. The New Humanitarian. https://www.irinnews.org/news/2019/02/05/un-palantir-deal-data-mining-protection-concerns-wfp

Perakslis, E. D. (2018). Using digital health to enable ethical health research in conflict and other humanitarian settings. *Conflict and Health*, *12*(1), 23. https://doi.org/10.1186/s13031-018-0163-z

Peters, M. D. J., Godfrey, C., Mcinerney, P., Baldini Soares, C., Khalil, H., & Parker, D. (2017). Chapter 11: Scoping reviews. In E. Aromataris & Z. Munn (Eds.), *Joana Briggs Institute Reviewer's Manual* (Issue September, pp. 6–24). https://reviewersmanual.joannabriggs.org/

Pham, P. N., & Vinck, P. (2012). Technology fusion and their implications for conflict early warning systems, public health, and human rights. *Health and Human Rights*, *14*(2), 106–117. http://www.jstor.org/stable/healhumarigh.14.2.106

Pictet, J. (1979). *The fundamental principles of the Red Cross*. http://www.ifrc.org/PageFiles/40669/Pictet Commentary.pdf

Raymond, N. A., Al Achkar, Z., Verhulst, S., & Berens, J. (2016). Building data responsibility into humanitarian action. *OCHA Policy and Studies Series*, *018*, 18. https://docs.unocha.org/sites/dms/Documents/TB18_Data Responsibility_Online.pdf

Raymond, N. A., & Card, B. (2015). *Applying Humanitarian Principles to Current Uses of Information Communication Technologies: Gaps in Doctrine and Challenges to Practice*. Harvard Humanitarian Initiative. https://hhi.harvard.edu/sites/default/files/publications/signal_program_humanitarian_princip

les_white_paper.pdf

Redmond, A. D. (2005). Needs assessment of humanitarian crises. *Bmj*, *330*(7503), 1320–1322. https://doi.org/10.1136/bmj.330.7503.1320

Russell, S. J., & Norvig, P. (2009). *Artificial Intelligence: A Modern Approach* (3rd editio). Prentice Hall.

Sandvik, K. B., Jumbert, M. G., Karlsrud, J., & Kaufmann, M. (2014). Humanitarian technology: a critical research agenda. *International Review of the Red Cross*, *96*(893), 219. https://doi.org/http://dx.doi.org/10.1017/S1816383114000344

Sarı, L., Hasegawa-Johnson, M., & Yoo, C. D. (2021). Counterfactually Fair Automatic Speech Recognition. *IEEE/ACM Transactions on Audio Speech and Language Processing*, *29*, 3515–3525. https://doi.org/10.1109/TASLP.2021.3126949

Savoldelli, A., Codagnone, C., & Misuraca, G. (2014). Understanding the e-government paradox: Learning from literature and practice on barriers to adoption. *Government Information Quarterly*, *31*(SUPPL.1), S63–S71. https://doi.org/10.1016/j.giq.2014.01.008

Scholte, J. A. (2014). Reinventing global democracy. *European Journal of International Relations*, *20*(1), 3–28. https://doi.org/10.1177/1354066111436237

Simms, B., & Trim, D. J. B. (2011). *Humanitarian Intervention: A History*. Cambridge University Press.

Singh, S. (2018). *Natural Language Processing for Information Extraction*. http://arxiv.org/abs/1807.02383

Slim, H. (2015). *Humanitarian Ethics: A Guide to the Morality of Aid in War and Disaster*. Oxford University Press.

Sphere Association. (2018). *The Sphere Handbook: Humanitarian Charter and Minimum*

Standards in Humanitarian Response (4th editio). Practical Action Publishing. https://doi.org/ISBN 978-1-908176-00-4

Spiegel, P. B., Burkle, F. M., Dey, C. C., & Salama, P. (2001). Developing public health indicators in complex emergency response. *Prehospital and Disaster Medicine*, *16*(4), 281–285. https://doi.org/10.1017/S1049023X00043430

Tansey, C. M., Pringle, J., Davé, A., Boulanger, R., & Hunt, M. (2018). Earthquakes to Floods: A Scoping Review of Health-related Disaster Research in Low- and Middle-income Countries. *PLoS Currents*. https://doi.org/10.1371/currents.dis.57d98a902a326361d88d54521e68b016

Tatman, R. (2017). Gender and Dialect Bias in YouTube's Automatic Captions. *Proc. of the First Workshop on Ethics in Natural Language Processing*, 53–59. http://www.aclweb.org/anthology/W17-1606

Tatoeba. (2022). *Tatoeba*.

Tegmark, M. (2017). *Life 3.0 : being human in the age of artificial intelligence*. Alfred A. Knopf.

The Engine Room and Oxfam. (2018). *Biometrics in the Humanitarian Sector*. https://www.theengineroom.org/wp-content/uploads/2018/03/Engine-Room-Oxfam-Biometrics-Review.pdf

Thylin, T., & Duarte, M. F. N. (2019). Leveraging blockchain technology in humanitarian settings - opportunities and risks for women and girls. *Gender and Development*, *27*(2), 317. https://doi.org/http://dx.doi.org/10.1080/13552074.2019.1627778

Translators Without Borders. (2018a). *The words between us: How well do enumerators understand the terminology used in humanitarian surveys? A study from northeast Nigeria*. https://translatorswithoutborders.org/wp-

content/uploads/2018/12/TWB_Nigeria_EnumeratorComprehension_Nov2018-1.pdf

Translators Without Borders. (2018b, April 19). *Communications Dashboard: Internally Displaced People in Northeast Nigeria*. https://translatorswithoutborders.org/communications-dashboard-internally-displaced-people-in-north-east-nigeria/

Translators Without Borders. (2019). *Language mapping: Putting communication needs on the map*. https://translatorswithoutborders.org/language-mapping/

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D. J., Horsley, T., Weeks, L., Hempel, S., Akl, E. A., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., Wilson, M. G., Garritty, C., … Straus, S. E. (2018). PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. *Annals of Internal Medicine, 169*(7), 467. https://doi.org/10.7326/M18-0850

UN Global Pulse. (2016). *Big Data for Development and Humanitarian Action: Towards Responsible Governance: Global Pulse Privacy Advisory Group Meetings*. http://unglobalpulse.org/sites/default/files/Big_Data_for_Development_and_Humanitarian_Action_Report_Final_0.pdf

UN OCHA. (2003). *Glossary of Humanitarian Terms*. OCHA. https://www.humanitarianresponse.info/sites/www.humanitarianresponse.info/files/documents/files/ocha glossary.pdf

UNHCR. (2006). *Master Glossary of Terms Rev. 1 United Nations High Commissioner for Refugees*. Geneva: UN High Commissioner for Refugees (UNHCR). http://www.refworld.org/docid/42ce7d444.html

UNHCR. (2017a). *Needs Assessment Handbook*.

http://needsassessment.unhcr.org/media/downloads/2017-05_NAH_BOOK_NEW.pdf

UNHCR. (2017b). *Standard WASH KAP Questionnaire*. http://wash.unhcr.org/download/wash-kap-survey-module-0-general/

United Nations. (2016). *One humanity: shared responsibility. Report of the Secretary-General for the World Humanitarian Summit*. New York: UN General Assembly. https://documents-dds-ny.un.org/doc/UNDOC/GEN/N16/025/68/PDF/N1602568.pdf?OpenElement

United Nations. (2022). *Report of the UN Secretary-General on the Protection of Civilians in Armed Conflict (S/2022/381)*. United Nations. https://documents-dds-ny.un.org/doc/UNDOC/GEN/N22/321/73/PDF/N2232173.pdf?OpenElement

van Wynsberghe, A., & Comes, T. (2020). Drones in humanitarian contexts, robot ethics, and the human–robot interaction. *Ethics and Information Technology*, *22*(1), 43–53. https://doi.org/10.1007/s10676-019-09514-1

Vannini, S., Gomez, R., Lopez, D., Mora, S., Morrison, J. C., Tanner, J., Youkhana, L., Vergara, G., & del Mar Moreno Tafurt, M. (2020). Humanitarian organizations' information practices: Procedures and privacy concerns for serving the undocumented. *Electronic Journal of Information Systems in Developing Countries*, *86*(1), 1–8. https://doi.org/10.1002/isd2.12109

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). *Attention Is All You Need*. https://arxiv.org/pdf/1706.03762.pdf

Wang, N., Christen, M., & Hunt, M. (2021). Ethical Considerations Associated with "Humanitarian Drones": A Scoping Literature Review. *Science and Engineering Ethics*, *27*(4), 1–21. https://doi.org/10.1007/s11948-021-00327-4

Weiss, R. J., Chorowski, J., Jaitly, N., Wu, Y., & Chen, Z. (2017). Sequence-to-sequence models

can directly translate foreign speech. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, *2017-Augus*, 2625–2629. https://doi.org/10.21437/Interspeech.2017-503

WHO. (2007). *Risk reduction and emergency preparedness*. World Health Organization. https://apps.who.int/iris/bitstream/handle/10665/43736/9789241595896_eng.pdf

WHS. (2016). *The Grand Bargain – A Shared Commitment to Better Serve People in Need*. Istanbul: World Humanitarian Summit. http://reliefweb.int/sites/reliefweb.int/files/resources/Grand_Bargain_final_22_May_FINAL-2.pdf

Yeophantong, P. (2014). Understanding Humanitarian Action in East and Southeast Asia: A Historical Perspective. In *HPG Working Paper* (Issue February). ODI/HPG. http://www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/8807.pdf

Chapter 2 Theoretical Frameworks

This dissertation follows a mixed-methods approach grounded in a pragmatist philosophy. This chapter will, first, explain this research paradigm underlying this study. Second, it will outline the values that have guided this research from the fields of humanitarian assistance, bioethics, and design science, respectively. Third, it will outline the methods chosen for this study.

2.1. Research Paradigm

A research paradigm, also referred to as a "worldview" (Creswell, 2018, p. 3), is understood here as the "basic set of beliefs that guide action" (Guba, 1990, cited in Creswell, 2018, p. 4). This paradigm is oftentimes described in terms of one's ontology, referring to a researcher's view about the nature of the world, or as Furlong and Marsh (2010, p. 18) have described it, "whether there is a 'real' world 'out there' that is independent of our knowledge of it." This study is rooted in the ontological stance that this is not an either-or question, that there is both an "external world independent of the mind as well as that lodged in the mind" (Creswell, 2018, p. 9) and that research should inform, or even participate in finding, solutions to real-world problems.

These sets of assumptions have led to the choice of a pragmatist research paradigm for this study. Pragmatism is a philosophy of research "in which the meaning of actions and beliefs is found in their consequences" (Morgan, 2014, p. 26). Moreover, pragmatism is deeply interested in the "intersubjective interactions between people and their environment" (McCaslin, 2012, p. 4), which requires taking into account all objects, subjects and their interactions. Pragmatist research means conducting inquiry for the sake of creating solutions to specific

problems, and choosing methods that contribute to this goal. As a result, pragmatist research follows a flexible inquiry process in which a research problem drives an ongoing iterative process of choosing and applying potential research designs and methods, and during which reflections on the potential consequences of each design and method inform the subsequent iteration and final design, or even serve to reformulate the initial problem (Morgan, 2014, pp. 27–34).

Researchers using a pragmatist worldview consider all possible research paradigms—as well as all potential methods—to be legitimate, as long as they are found to serve a given research objective. Other paradigms, such as positivism, post-positivism, critical theory, and constructivism, are seen by some of their adherents as indelible categories that mark each researcher and that necessarily lead to specific sets of potential methods that can be considered. Furlong and Marsh (2010, p. 17) have likened research paradigms (specifically ontologies and epistemologies that are used to distinguish different paradigms) as a "skin," which "cannot be put on and taken off whenever the researcher sees fit." As such, many theorists have historically viewed some different ontologies and methodological approaches as "incommensurate" or "contradictory and mutually exclusive" (Guba & Lincoln, 2005, p. 117). Pragmatism, on the other hand, has "no interests in dualistic debates around philosophical concepts" (Biesenthal, 2014, p. 649). Pragmatism is often compared to—and borrows from—constructivism when humans and cultural practices are involved (see, for example, Hickman et al., 2009; Neubert, 2001). It therefore emphasizes the epistemological view that much of the world is socially constructed and all knowledge is social knowledge based on experience. It also allows—as would post-positivists—that reality exists apart from human experience. As Morgan (2014, p. 40) points out, some researchers who adopt pragmatism see metaphysical separation as a useful

way to understand social science research and believe instead that applying such different worldviews can be useful, depending on the research in question.

Axiology, the branch of philosophy that focuses on values (among other issues), forms a crucial part of the "basic foundational philosophical dimensions of [the] paradigm proposal" (Guba & Lincoln, 2005, p. 200). In pragmatism, values play a central role in informing both the motivation behind the research as well as for understanding the values held by research subjects themselves (McCaslin, 2012). Based on work by The Ethics Centre (2020), this study defines values as ends in themselves which ought to be respected and realized in practice. Principles are understood here as the specific means with which one can order and accomplish these values. This study does not use the term "moral" as it can practically be understood to be synonymous with "ethical" (Singer, 2021), although varying distinctions exist in the philosophical and legal spheres between these terms (Hazard, 1995).

The following section describes the values that underpin the different disciplines that intersect in this study. This section is separated into three parts: humanitarian values, bioethical values, and design values.

2.2. Humanitarian Values

It is commonly understood that ethical values and principles are essential in the provision of humanitarian assistance where an individual's or an organization's choices can be far-reaching in determining the well-being and health outcomes of other people or even large populations. The ethical values of compassion and egalitarianism that motivate many people to assist others in need are typically seen as rooted in human history (see, for example, Barnett, 2011, p. 8; Suzman, 2017; Wade, 2006). Likewise, these values have also been prominent precursors to humanitarian assistance across different parts of the world—such as, for example, in East and

Southeast Asia (Yeophantong, 2014), China (Krebs, 2014), the Middle East (Moussa, 2014), or Latin America (L. L. Martínez, 2016).

Modern humanitarian assistance is made up of a highly fragmented patchwork of actors performing under a loose and complex web of governance structures, and with limited control exercised by governments (Barnett, 2013). As a result, a single set of ethical values that is universally accepted by or even mandated for all organizations providing humanitarian assistance does not (yet) exist. Instead, as discussed below, there are several overlapping humanitarian ethical values frameworks that serve as foundational documents that were intended to ensure that humanitarian professionals make ethical decisions in their work. In addition, international humanitarian law and human rights law provide some legal obligations primarily specifying obligations of states or warring parties during conflicts, the rights of civilians, and the protected status of organizations providing assistance (Darcy, 2004; Leaning, 1999).

The first declaration to better guide the behavior of humanitarian organizations and their staff was established in 1965 with the Fundamental Principles of the Red Cross (IFRC, 2006; Pictet, 1979), hereinafter referred to as the *Fundamental Principles*. The first four principles in this document—humanity, impartiality, independence, and neutrality—have since become part of international law through two resolutions passed by the United Nations General Assembly (UNGA) (OCHA, 2012) and are commonly referred to as the *Humanitarian Principles*. Definitions of each of the *Humanitarian Principles* are shown in Table 1. The original author of the *Fundamental Principles*, Jean Pictet (1979, p. 8), understood principles as “a rule, based upon judgement and experience, which is adopted by a community to guide its conduct.” The *Humanitarian Principles* are now widely cited among many humanitarian organizations in official documents as high-level guidance for their work (see, for example, MSF, n.d.).

Table 1

Definitions of the Humanitarian Principles (OCHA, 2012)

Principle	Definition
Humanity	Human suffering must be addressed wherever it is found. The purpose of humanitarian action is to protect life and health and ensure respect for human beings.
Impartiality	Humanitarian action must be carried out on the basis of need alone, giving priority to the most urgent cases of distress and making no distinctions on the basis of nationality, race, gender, religious belief, class or political opinions.
Neutrality	Humanitarian actors must not take sides in hostilities or engage in controversies of a political, racial, religious or ideological nature.
Independence	Humanitarian action must be autonomous from the political, economic, military or other objectives that any actor may hold with regard to areas where humanitarian action is being implemented.

During the 1990s, three additional sets of ethical principles were established by leading international humanitarian organizations with the aim of addressing perceived failures in a number of humanitarian crises during that time (see, for example, Anderson, 1999; de Waal, 1997; Orbinski, 2009). This included, first, the *Code of Conduct for the International Red Cross and Red Crescent Movement and NGOs in Disaster Relief* (IFRC, 1994, "Code of Conduct"), and, second, the *Humanitarian Charter* and the *Protection Principles* that were first published in 1998 as part of the Sphere Guidelines (Sphere Association, 2018; Walker & Purdin, 2004).

During the 2000s, four additional initiatives by networks of humanitarian organizations attempted to complement these general principles into more concrete rules aiming to increase accountability and effectiveness. These included Sphere's (2004) *Core Standards*, People in Aid's (2003) *Code of Good Practice*, the Humanitarian Accountability Partnership's (2010) *Standard in Accountability and Quality Management*, and Groupe URD's (2018) *Quality COMPAS* framework. The *Core Humanitarian Standard on Quality and Accountability* (CHS)

then replaced or integrated these documents, along with several others (CHS, 2014), following several years of consultations. An overview of the six documents discussed in this section can be found in Table 2, including the number of articulated principles included in each one.

Table 2

Selection of Ethical Frameworks in the Humanitarian Assistance Sector

Framework	Author	Published	Principles
Fundamental Principles	IFRC	1965, revised 1986	7
Humanitarian Principles	UNGA OCHA	1991 and 2004 2012	4
Code of Conduct for the International Red Cross and Red Crescent Movement and NGOs in Disaster Relief (Code of Conduct)	IFRC & ICRC	1994	10 (including 4 Humanitarian Principles)
Humanitarian Charter	Sphere Project	2000, rev 2004-2018	11 (including 6 rights or beliefs)
Protection Principles	Sphere Project	2000, revised 2004-2018	4
Core Humanitarian Standard on Quality and Accountability (CHS)	CHS Alliance	2014	9

Despite the proliferation of these documents with their additional principles, the four *Humanitarian Principles* are either included explicitly (as, for example, in the *Code of Conduct* or the *CHS*) or included indirectly as an "integral part" (Sphere Association, 2018, p. 5), as in the case of the *Humanitarian Charter*. As a result, these principles often form part of induction training for new humanitarian staff, and are included in lectures on the topic of humanitarian assistance. As Gordon and Donini (2015, p. 79) point out, by elevating them in this fashion, organizations established them "as global, permanent and immutable talismans of access, and as central motifs *qua* objectives of the humanitarian discourse." Pragmatically, this means that for

humanitarian professionals working at international organizations, governments, and non-governmental organizations, these four principles have in practice become enshrined as values worthy of achieving in their own right.

Previous studies have shown how difficult it can be to apply the *Humanitarian Principles* in everyday practice (Hilhorst & Schmiemann, 2002); to guide the use of information technology (Raymond & Card, 2015), or to map humanitarian organizations' ethical obligations (Broussard et al., 2019). As shown by Gordon and Donini (2015), some humanitarian organizations and donors have therefore questioned the relevance of the independence and neutrality principles in novel situations and conflicts, leading to a "new humanitarianism" that may result in a new set of principles that aim to responds to changing circumstances or that may leave the neutrality principle to be rendered optional (Van Mierop, 2015; Walker, 2005). Similarly, Orbinski (1994, p. 7) has suggested that the principles of independence and neutrality can be best understood as mechanisms of the humanity principle (often referred to as the humanitarian imperative) that should be used as necessary to reach and provide benefits to the greatest portion of affected populations in conflict settings.

Several attempts to analyze contemporary humanitarian ethical values and principles have been made. Slim (2015) created the most extensive overview of relevant documents and their described principles and how they can be applied together, although his work preceded the finalization of the *CHS* which made some of the suggested new principles arguably obsolete. Thorley and Henrion (2019, p. 3) reviewed documents used by humanitarian and development organizations for guiding research and evaluation activities, finding that "guidelines do not consistently provide clear definitions of the ethical principles underpinning them." Broussard (2019) reviewed how health-related ethical challenges documented in the literature mapped onto

the four humanitarian principles from 1965 and established eight "ethical obligations." The Harvard Humanitarian Initiative's Signal Code (Greenwood et al., 2017) provides additional analysis of humanitarian ethical obligations related to data processing by also applying human rights law.

This dissertation follows the proposal by Slim (2015, p. 62) that considers humanity to be understood as an absolute value, whereas impartiality, independence, and neutrality are understood as necessary principles that help to achieve this value. This view is what Gordon and Donini (2015, p. 83) refer to as "traditional or classical humanitarianism." Although he named it as one of the *Fundamental Principles*, Jean Pictet (1979, p. 8) seems to agree with this view. Indeed, he wrote that the humanity principle embodied the "profound motivation of the Red Cross, from which all the other principles are derived." Identifying an absolute value can be crucial as a means to resolve conflicts between principles and to help make an *ethical* decision. For example, Médecins Sans Frontières decided to speak out and even request military actors to intervene given the extreme conditions during the Rwandan genocide—even if this meant not being perceived as neutral—in order to achieve the humanitarian imperative (Orbinski, 2009).

However, the exact content and meaning of humanity (or the humanitarian imperative) varies from document to document as each tried to improve on the wording rather than adopting an existing definition. A detailed review of the humanity value (though often referred to as a principle) of the six foundational texts cited above shows some important differences (see Table 3). Even though most documents restate the four *Humanitarian Principles*, their authors applied sometimes significant wording differences. For example, the importance of activities related to safeguarding human health and saving lives do not appear in the Code of Conduct version of *humanity*—nor in the Humanitarian Charter or the Protection Principles. The wording in the

UNGA resolution that enshrined the *Humanitarian Principles* into international law says that human suffering should be “addressed” (OCHA, 2012, p. 1) rather than to “prevent and alleviate” it, as specified in the *Fundamental Principles* (IFRC, 2006, p. 5).

Based on this analysis, four overlapping themes were identified (displayed with different color coding in Table 3), which this study uses as a working definition of the value of humanity, and therefore as the purpose of humanitarian assistance: (1) preventing and alleviating human suffering, (2) providing protection and safeguarding of human dignity, (3) safeguarding human health, and (4) conducting these activities anywhere in the world where they may be needed.

Table 3

Varying Definitions of the Humanity Value Used in Different Foundational Texts, With Overlapping Themes Highlighted

Document	Definition text
IFRC Fundamental Principals	"The International Red Cross and Red Crescent Movement, born of a desire to bring assistance without discrimination to the wounded on the battlefield, endeavours, in its international and national capacity, to prevent and alleviate human suffering wherever it may be found. Its purpose is to protect life and health and to ensure respect for the human being. It promotes mutual understanding, friendship, cooperation and lasting peace amongst all peoples." (IFRC, 2006, p. 5)
Humanitarian Principles	"Human suffering must be addressed wherever it is found. The purpose of humanitarian action is to protect life and health and ensure respect for human beings." (OCHA, 2012, p. 1)
Code of Conduct	"The humanitarian imperative comes first: The right to receive humanitarian assistance, and to offer it, is a fundamental humanitarian principle which should be enjoyed by all citizens of all countries. As members of the international community, we recognise our obligation to provide humanitarian assistance wherever it is needed. Hence the need for unimpeded access to affected populations is of fundamental importance in exercising that responsibility. The prime motivation of our response to disaster is to alleviate human suffering amongst those least able to withstand the stress caused by disaster. When we give humanitarian aid it is not a partisan or political act and should not be viewed as such." (IFRC, 1994, p. 3)
Humanitarian Charter	"The Humanitarian Charter expresses our shared conviction as humanitarian agencies that all people affected by disaster or conflict have a right to receive protection and assistance to ensure the basic conditions for life with dignity. We believe that the principles described in this Humanitarian Charter are universal, applying to all those affected by disaster or conflict, wherever they may be, and to all those who seek to assist them or provide for their security. These principles are reflected in international law, but derive their force ultimately from the fundamental moral principle of humanity: that all human beings are born free and equal in dignity and rights. Based on this principle, we affirm the primacy of the humanitarian imperative: that action should be taken to prevent or alleviate human suffering arising out of disaster or conflict, and that nothing should override this principle." (Sphere Association, 2018, p. 28)
Protection Principles	"[…] ensure that the rights of affected persons and the obligations of duty bearers under international law are understood, respected, protected and fulfilled without discrimination. Protection is about taking action to keep people safe from violence, coercion and deliberate deprivation." (Sphere Association, 2018, p. 36)
Core Humanitarian Standard (Section V)	"The primary motivation of any response to crisis is to save lives, alleviate human suffering and to support the right to life with dignity. Humanitarian organisations recognise that the humanitarian imperative comes first and seek, therefore, to provide humanitarian assistance wherever it is needed." (CHS, 2014, p. 8)
	Legend of overlapping themes highlighted in texts: (1) Preventing and alleviating human suffering (2) Providing protection and safeguarding of human dignity (3) Safeguarding human health (4) Conducting these activities anywhere in the world where they may be needed

Whereas the value of *humanity* appears to be widely supported by humanitarian organizations, as evident in the various documents discussed above, the lack of a shared definition can be a challenge for studies related to ethical issues in this field. As a result, some humanitarian practitioners and researchers have looked to the fields of bioethics and research ethics for a more specific and inclusive set of values, as discussed in the following section.

2.3. Bioethical Values

The ethical value categories of *autonomy*, *beneficence*, *non-maleficence*, and *justice*, are widely used in the fields of bioethics and research ethics, and have been defined and discussed in their application in detail by Beauchamp and Childress (2019) in multiple volumes dating from 1985. A definition of each ethical value category is provided in Table 4. The four categories reflect work largely in the decades preceding 1985 that have aimed to better protect research participants, including the Nuremberg Code (*Trials of War Criminals before the Nuernberg Military Tribunals under Control Council Law No. 10 (Volume 2)*, 1949), the 1964 Declaration of Helsinki (World Medical Association, 2013) and the Belmont Report (National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, 1978).

Table 4

Definitions of Each Ethical Value Category, Based on Beauchamp and Childress (2019, p. 16)

Principle	Definition
Respect for autonomy	Respecting the decision-making capacities of autonomous persons
Beneficence	Providing benefits and balancing benefits against risks and costs
Non-maleficence	Avoiding the causation of harm
Justice	Distributing benefits, risks, and costs fairly

Several studies focused on humanitarian assistance have also adopted these value categories as better operational ethical categories for analyzing ethical issues (see, for example, Cawthorne & Wynsberghe, 2019; Pham & Vinck, 2012). Similarly, nurses and doctors working at humanitarian medical organizations are expected to rely on established ethical codices, which are often set by medical associations in their home countries, and by the World Medical Association (Parsa-Parsi, 2022). Many humanitarian practitioners also have a medical or graduate training in which these categories are often used in lectures on ethical practices (see, for example, Barman et al., 2020; Marco et al., 2011). For health practitioners, adhering to ethical codes encompassing these value categories (or overlapping with them) is typically mandatory to be allowed to practice, and these codes are also expected to be applied in humanitarian crisis settings (Schwartz et al., 2012). For example, the *Ethical Principles of Health Care in Times of Armed Conflict and Other Emergencies* (ICRC, 2015, p. 1) proscribe that ethical codes to be used in conflict zones "are the same as the ethical principles of health care in times of peace." However, as Hunt (2008) explains, applying these codes in humanitarian settings can sometimes lead to tensions between ethical values. In academia, university ethical review boards generally require a priori approval for research involving human subjects or animals as a way of implementing ethical codes. For research purposes, several humanitarian organizations have mirrored such structures, and created similar review boards, and even created specific innovations such as pre-approved research protocols, for example in the case of Médecins Sans Frontières (Benelli & Low, 2019; Eckenwiler et al., 2015; MSF Ethics Review Board, 2013; Schopper et al., 2015).

This study considers the four value categories of autonomy, beneficence, non-maleficence, and justice, using the definitions shown above, as providing additional ethical

guidance for humanitarian assistance in general, and this research in particular. A closer examination of these four ethical value categories and existing principles from the humanitarian ethical codes described in the previous section shows that a considerable overlap already exists.

One way of mapping humanitarian ethical principles to the bioethical value categories is shown in Table 5. This overlap is discussed in more detail below.

Table 5

Humanitarian Ethical Principles Mapped to Bioethical Value Categories

Value Categories	Humanitarian Principals	Code of Conduct	Humanitarian Charter	Protection Principles	Core Humanitarian Standard
Respect for autonomy	1. Humanity	5. Respect culture and custom 6. We shall attempt to build disaster response on local capacities 7. Ways shall be found to involve programme beneficiaries in the management of relief aid 9. We hold ourselves accountable to both those we seek to assist and those from whom we accept resources	8. […] the affected population is at the centre of humanitarian action, and recognise that their active participation is essential […]	4. Help people to claim their rights	4. Humanitarian response is based on communication, participation and feedback. 8. Staff are supported to do their job effectively, and are treated fairly and equitably.
Beneficence	1. Humanity 3. Independence 4. Neutrality	8. Relief aid must strive to reduce future vulnerabilities to disaster as well as meeting basic needs	11. […] adopt the Core Humanitarian Standard and the Minimum Standards as accepted norms […] 12. […]we undertake to make our responses more effective, appropriate and accountable […]	3. Assist people to recover from the physical and psycho- logical effects of threatened or actual violence, coercion or deliberate deprivation	2. Humanitarian response is effective and timely 6. Humanitarian response is coordinated and complementary. 7. Humanitarian actors continuously learn and improve. 9. Resources are managed and used responsibly for their intended purpose.
Non-maleficence	1. Humanity	10. In our information, publicity and advertising activities, we shall recognise disaster victims as dignified humans, not hopeless objects	9. […] we aim to minimise any negative effects of humanitarian action on the local community […]	1. Enhance people's safety, dignity and rights and avoid exposing them to further harm	3. Humanitarian response strengthens local capacities and avoids negative effects 5. Complaints are welcomed and addressed.
Justice	1. Humanity 2. Impartiality		6. […] the principle of impartiality, which requires that it be provided solely on the basis of need and in proportion to need. This reflects the wider principle of non-discrimination: that no one should be discriminated against on any grounds […]	2. Ensure people's access to impartial assistance, according to need and without discrimination	1. Humanitarian response is appropriate and relevant.

2.3.1. Respect for autonomy

Respecting the autonomous decision-making capacity of individuals or communities has been a major factor in various humanitarian reform efforts since at least the 1990s. Starting with the *Code of Conduct* in 1994, new principles were established that attempted to remedy a perceived lack of autonomy. A subtle shift can be observed over time as different principles were elaborated to reflect changing norms regarding the role that affected people can and should play in deciding on the most relevant assistance. The 1994 *Code of Conduct*'s used careful wording, proposing that ways "shall be found to involve programme beneficiaries" (IFRC, 1994, p. 4). In contrast, the *CHS* (2014, p. 9) took a more forceful stance 20 years later with a "quality criterion" that requires that "humanitarian response is based on communication, participation and feedback." One of the most well-known examples in research and biomedical ethics is the requirement of informed consent. In the humanitarian assistance context, informed consent for data collection is now a requirement for many organizations, especially when data is collected for the express purpose of research (International Rescue Committee, 2018).

2.3.2. Beneficence

Beneficence (providing benefits and balancing them against risks and costs) for affected people and communities is the stated goal of all organizations involved in humanitarian assistance, as shown in the consistent reference to the prevention or alleviation of human suffering in all the documents discussed in the previous section. Yet, in the response to conflicts in the 1990s, stories abounded about organizations that conducted activities that lacked tangible benefits for people affected by humanitarian crises, whether due to wrong assumptions, lack of training, lack of oversight, trying to impress donors—or a combination of these factors (see, for example, Polman, 2010; Terry, 2002). These episodes spurred the re-articulation and

contemporizing of ethical principles that pledged goals that hitherto were taken for granted. As a result, explicit goals that can be grouped under the value category of beneficence can be found, for example, in the *Humanitarian Charter* which obliges humanitarian actors "to make our responses more effective, appropriate and accountable" (Sphere Association, 2018, p. 32), or in the *CHS* (2014, p. 9), which requires that funding should be "managed and used responsibly for their intended purpose."

2.3.3. Non-maleficence

Non-maleficence as an ethical value category (avoiding the causation of harm) was challenged by events in the 1990s, which led to the growing recognition that humanitarian assistance can inadvertently lead to harmful consequences for affected populations, such as by potentially prolonging conflict or increasing the risk of violence (see, for example, Anderson, 1999; de Waal, 1997; Lischer, 2003; J. C. Martínez & Eng, 2016; Narang, 2015; Orbinski, 2009). This, in turn, has correspondingly led to an increase in norms that directly relate to non-maleficence, which has often been reinterpreted as the maxim "first, do no harm." Contrary to the earlier *Principles* and the *Code of Conduct*, the *Humanitarian Charter* requested that humanitarian assistance should also "minimise any negative effects of humanitarian action," such as rendering "civilians more vulnerable to attack" (Sphere Association, 2018, p. 31).

The value of non-maleficence is by now also widely integrated in more specific guidelines related to processing sensitive information. Recognizing the military value that program information and data can have, the *Protection Principles* also implores that "people are not put at risk as a result of the way that humanitarian actors record and share information" and requests that policies are in place to ensure "the safety of survivors and of staff" (Sphere Association, 2018, p. 40). The *CHS* requires that affected people are not negatively affected by

humanitarian action, and urges organizations to put systems in place to "safeguard any personal information collected from communities and people affected by crisis that could put them at risk" (CHS, 2014, p. 12). The 2018 edition of the *Sphere Handbook* expands on this last point by explaining that "clear and comprehensive policies on data protection" should be "aligned with international standards and local data protection laws" (Sphere Association, 2018, p. 62).

2.3.4. Justice

The ethical value category of justice (distributing benefits, risks, and costs fairly) in many ways overlaps with the humanitarian principle of impartiality. Jean Pictet commented later that impartiality should, in fact, have been called "equality" in the *Fundamental Principles*: "While it was not particularly appropriate to have classified the principles of non-discrimination and of proportionality under the same heading, it was incorrect to have given this heading the designation of Impartiality, for this is a personal quality of an individual called upon to make a judgment or choice, or for 'the man from the Red Cross' to distribute relief or give care." (Pictet, 1979, p. 31) As a result of Pictet's explanations, impartiality can be understood to include three related but separate sub-principles that require that assistance be given based on 1) proportionality (based on relative need), 2) non-discrimination (irrespective of ethnic, class, gender, or other attributes), and 3) impartiality (irrespective of personal preferences). For example, organizations such as Médecins Sans Frontières have defined impartiality to include both the proportionality and non-discrimination criteria, explaining that "those in the most serious and immediate danger will receive priority" (Binet & Saulnier, 2019, p. 158).

Whereas a strict reading of the term impartiality would only be concerned with the equal distribution of benefits, the inclusion of proportionality and non-discrimination in the definition

show the significant overlap with the ethical value category of Justice (defined above as "distributing benefits, risks, and costs fairly"). Given the limitations of organizations and extraordinary needs in many humanitarian crises, decisions of which assistance to provide, where and to whom, often require hard trade-offs. In such situations, the objective of fairly distributing benefits, risks, and costs is pursued in practice—even if humanitarians may call this impartiality rather than justice.

It is beyond the scope of this study to determine whether these four value categories will find broader adoption by humanitarian organizations—especially in conjunction with a more clearly articulated version of the value of humanity. However, two facilitating factors may favor a more widespread adoption. First, many humanitarian practitioners have previously been trained in research or medical ethics and are therefore quite familiar with these categories. An increasing overlap between the epistemic communities of academic research, healthcare, and humanitarian assistance may therefore make it easier to embrace and engage with the four value categories. Second, adopting an existing list of values also prevents the difficult task of establishing a completely new set of ethical values that would require significant time and resource for consultations between humanitarian professionals that may, in the end, result in a list that appears very similar.

Apart from ethical values, pragmatism also requires investigating and documenting aesthetic values. These are discussed in the following section.

2.4. Design Values

Davis (1989) argued that software needs to be designed in a way that is both useful and easy to use, a concept also known as usability, and this supports the position taken in this study. This goal is encapsulated in ISO standard 9241-11:2018 (International Organization for

Standardization, 2018) which defines usability as the "extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use." One widely used approach in web and software design is user-centered design (UCD), which is a framework for flexibly and iteratively creating innovations that are supported by research into user realities as well as the institutional context in which innovations are intended to be used (Holden & Boustani, 2021). As described first by Gould and Lewis (1983), UCD typically requires 1) a thorough understanding of the eventual users in order to take their needs and environment into account, 2) empirical evaluations of user interactions, and 3) iterative cycles of designing and building new systems or services. Empirical evaluations of usability, often referred to as user testing, is widely seen as an essential step that should be employed at different intervals during the design process (rather than a single evaluation at the end). However, implementing this step in practice is often beset with many practical challenges that need to be addressed through proper research design (Cornet et al., 2020).

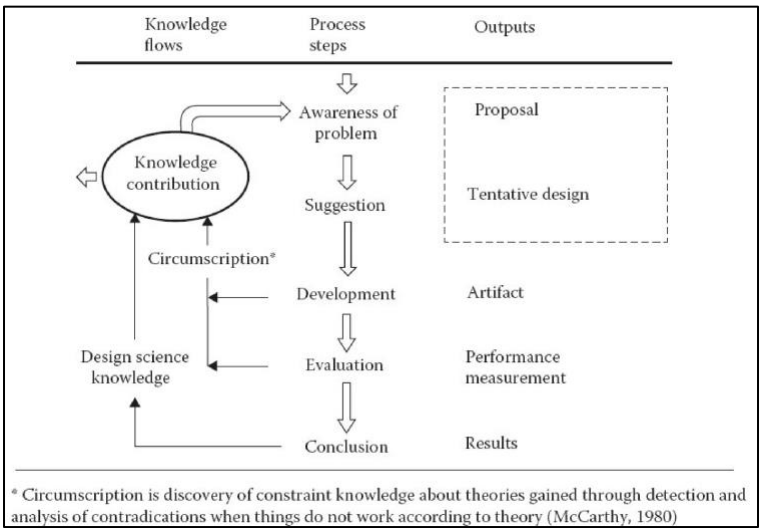
Proper application of this method means, by definition, an openness to change: If an evaluation shows that a particular feature is useless without significant changes, good UCD allows that the final innovation may differ from what was initially discussed and what a client may have expected. Innovations that were designed without proper UCD methods can render new ideas useless when they are rejected by organizations or professionals due to poor usability, or even have the potential to cause harm (Cornet et al., 2019). Publicly-funded software projects are more likely to be affected by this issue (Anthopoulos et al., 2016; Savoldelli et al., 2014), the consequences of which can be particularly acute in critical settings such as during the provision of humanitarian assistance or healthcare (Ben-Zeev et al., 2015; Kushniruk & Nøhr, 2016).

Design Science Research (DSR) is a "is a body of knowledge about artificial (man-made) objects and phenomena designed to meet certain desired goals" (Simon, 1996, p. 1). As an academic discipline it is therefore well suited to a pragmatic approach of studying the conditions for, as well as the process of creating new software. In this regard, DSR is closely related and overlaps with action research, both of which are "pursued with an explicit intention of achieving desired social or organizational ends" (Stewart, 2014, p. 247). Simon (1996) originally referred to it as *sciences of the artificial*, which he defined to include natural phenomena as well as the study of socially constructed institutions.

DSR underlines the tenet that in order to create new technologies to address a social problem, proper design goes far beyond trial and error. Instead, "design work entails disciplined, reflective discovery, and development of concepts for seeing what is possible and methods for realizing what is possible" (Jackson & Aakhus, 2014, p. 3). The DSR cycle, shown in Figure 3, typically starts with an awareness of a problem and a suggestion for a solution, resulting in a proposal and a tentative design. The further design and development of the new solution (also known as "artifact") then leads to various forms of evaluation, leading to iterative changes to the initial problem, suggested solutions, and new development—until a conclusion can be reached at the end of the research project (Vaishnavi & Kuechler, 2015).

Figure 3

Design Science Research Cycle, by Vaishnavi and Kuechler (2015, p. 14)



In DSR, the contribution to scientific knowledge takes place at every step of the aforementioned design cycle. As Vaishnavi & Kuechler (2015) explain, DSR generates knowledge through a complete awareness of the problem based on initial research and through experience gained from designing and evaluating the pilot tool. Specific design science knowledge can then be used to allow generalization of the work in other situations.

This study considers the concept of usability, as defined above, as a high-level value that is used to guide both the research and the creation of new features using Natural Language Processing (NLP) for the KoboToolbox software for humanitarian needs assessment that forms the core of this project. It considers the steps of the DSR cycle as well as best practices of the UCD framework as crucial and complementary steps for implementing this value.

The pragmatist research paradigm and the humanitarian values described in the previous sections undergird this study and have served as a guide for selecting the specific research methods. Those methods are summarized in the following section.

2.5. Methods

This study uses a mixed-methods approach by using key informant (KI) interviews, a scoping review of the literature, and by collecting observations of, and feedback from test users about the usability of newly designed NLP features for the KoboToolbox software for use in humanitarian needs assessments. These methods are described in greater detail in each of the empirical chapters of this dissertation (Chapters 4, 5, and 6).

2.5.1. Key Informant Interviews

The study selected 23 KIs using exponential discriminative snowball sampling of international professional staff with prior experience collecting data for humanitarian needs assessments (HNA) at a wide array of humanitarian organizations. Additional respondents were sought to achieve maximum heterogeneity of sampled organizations, including United Nations (UN) agencies, the Red Cross/Red Crescent Movement, as well as nongovernmental organizations (NGOs).

A semi-structured interview guide was developed that included questions about the contextual considerations for using quantitative and qualitative methods in HNA, the constraints related to qualitative interview methods, and potential solutions for improving the usefulness of qualitative interview methods in HNA. Data collection, including an initial pilot with three respondents, took place between November 2020 and May 2021. Interviews were conducted by the author using the Zoom platform (Gray et al., 2020), while audio was recorded for verbatim transcription.

Transcripts were analyzed using MAXQDA 2020 software (*MAXQDA 2020*, 2019) using an iterative approach consistent with stablished methods (Belgrave & Seide, 2018; Chun Tie et al., 2019; Corbin & Strauss, 2015). After applying initial codes through an open coding

approach, additional codes were developed through constant comparative analysis of previously-coded transcript segments. Themes were categorized according to whether they related to issues that could be addressed through technological means or not. This study follows O'Brien et al.'s (2014) Standards for Reporting Qualitative Research checklist. Full methodological details are provided in Chapter 4.

2.5.2. Scoping Review of the Literature

A scoping review was conducted to map the range of ethical issues that have been raised in the academic literature regarding data processing of people affected by humanitarian crises. This method was chosen as it is best suited for generating a broad overview of relevant evidence; for examining emerging areas of research; for clarifying key concepts, and for identifying gaps in the literature (Peters et al., 2015). This study followed the scoping review method first described by Arksey and O'Malley (2005) as further refined by Levac et al. (2010) and follows the framework maintained by the Joanna Briggs Institute (Peters et al., 2017). A detailed study protocol was developed prior to data collection and screening and was revised based on feedback received from dissertation committee members, and incorporating the results from a pilot conducted for this study November–December, 2019. The protocol follows the Preferred Reporting Items for Systematic Reviews and Meta-analysis for Protocols—Extension for Scoping Reviews, or PRISMA-ScR (Tricco et al., 2018). As shown in Chapter 5, specific eligibility criteria were established to define the parameters required to establish the search of academic databases. This included definitions and decisions on key words related to concepts that include ethical issues, humanitarian crises, and data processing for informing humanitarian assistance.

A comprehensive search of the literature was conducted using Ovid, Ebsco, Web of Science, and Proquest to search 20 databases for relevant studies. Keywords were selected and piloted in multiple iterations to identify all potentially relevant articles, using a sample of 34 previously-identified studies that were used as a minimum search target. After an initial search showed that only 13 were included, the database search was repeated over several iterations with additional terms until all 34 studies were reflected in the results. Significant efforts were made to establish these keywords, due to the lack of a shared nomenclature across disciplines that could be relied on for identifying studies relating to humanitarian assistance, ethical issues, and data processing.

Study selection, coding, and data extraction were done using the DistillerSR systematic review software (Evidence Partners, 2022). Using the *a priori* eligibility criteria, this study used questionnaires for selecting citations during discrete title, abstract, and full text review stages. Two reviewers independently selected studies during each screening stage. Data extraction forms were first piloted and refined based on discussions between reviewers. Results were summarized quantitatively (using frequencies) and qualitatively (using descriptive analytics). Ethical issues related to data processing that were entered in text form were coded using SPSS 25. Codes for ethical issues were updated by creating new codes based on new observations and through retrospective reviews of previously collected data. Further methodological details are provided in Chapter 5.

2.5.3. Design of New Software Features and User Testing

Based on the needs identified during the qualitative research with KIs, a proposed high-level architecture of new features for the KoboToolbox software was developed by the researcher to support the needs expressed by specialists involved in processing and analyzing

HNA data. This was then used to inform the iterative creation of comprehensive interactive prototypes. The first set of interactive prototypes focused on features to view and organize additional audio files in survey data, the ability to do transcriptions and translations—both manually and using automated means—as well as ways of modifying and managing the generated text for different languages. The second interactive prototypes focused on a qualitative analysis feature for coding open-ended interview responses. Software engineers at KoboToolbox then implemented the new features, with code published in an open-source repository (KoboToolbox, 2022). Designs as well as software code were updated to address important issues found during usability testing sessions and feedback received from participants (see below).

The research participants involved in user testing were 14 professional staff involved in humanitarian assistance who had prior experience working with transcription and translation tools or with analyzing qualitative data in general. They took part in two testing sessions that were intended to test the usability of the prototypes by systematically recording observations and oral feedback (Travis & Hodgson, 2019). Only participants who spoke either English, French, or Spanish were considered. Initial recruitment was done by posting a message on a community forum (https://community.kobotoolbox.org) to ask for volunteer participants with relevant experience, and also by specifically asking for volunteer participants at Translators Without Borders. For Round 1, six participants were recruited. For Round 2, eight participants were recruited. Data collection took place in October 2021 and in July 2022 for the first and second round of user testing, respectively. User testing was conducted remotely using the Zoom platform (Gray et al., 2020). During each testing session, participants were asked to look at various prototype designs or access the KoboToolbox software on their computer and interact

with particular features. They were given simple prompts to execute several tasks, such as “You would like to translate this transcript from Hausa to English using the automated option. How would you do this?” Participants were told that they could ask questions about functionality and were asked open-ended questions at the end of each Round about their overall experience.

Detailed anonymized notes were taken for each user testing session by the dissertation researcher. Notes about observations for each task as well as a summary of responses given to each open-ended debriefing question were entered into a spreadsheet during the user testing session and were further expanded or corrected based on the recording, as needed. An inductive summary of notes was created after each of the two rounds of user testing to extract relevant comments.

2.6. Conclusion

This chapter has, first, explained the choice of the pragmatist research paradigm for this study, as well as the relevant ontological and epistemological assumptions underlying this study. Second, it outlined the values that have guided this research from the fields of humanitarian assistance, bioethics, and design science, respectively. It also provided a proposed bridge between the bioethical value categories and the humanitarian principles by exploring and identifying commonalities of overlap. Third, it summarized the methods chosen for this study, which are explained in more detail in each of the empirical chapters.

The following chapter provides an overview of the design concept that inspired the overall study by describing how new technological innovations can be used to transcribe, translate, and analyze large sets of qualitative responses with a view to improving the quality and effectiveness of humanitarian assistance. It also describes some of the ethical challenges that

might arise from such a design concept and strategy in the use of technological innovations in humanitarian emergency settings.

2.7. References

Anderson, M. B. (1999). *Do no harm: How aid can support peace – Or war*. Lynne Rienner Publishers.

Anthopoulos, L., Reddick, C. G., Giannakidou, I., & Mavridis, N. (2016). Why e-government projects fail? An analysis of the Healthcare.gov website. *Government Information Quarterly*, *33*(1), 161–173. https://doi.org/10.1016/j.giq.2015.07.003

Arksey, H., & O'Malley, L. (2005). Scoping studies: Towards a methodological framework. *International Journal of Social Research Methodology*, *8*(1), 19–32. https://doi.org/10.1080/1364557032000119616

Barman, B., Srivastava, T., Sarma, A., & Nath, C. (2020). Effectiveness of formal training in bioethics of 3 rd semester undergraduate medical students in recognizing bioethical issues and principles in patient care. *Journal of Family Medicine and Primary Care*, *9*(6), 2871. https://doi.org/10.4103/jfmpc.jfmpc_405_20

Barnett, M. N. (2011). *Empire of humanity: A history of humanitarianism*. Cornell University Press.

Barnett, M. N. (2013). Humanitarian Governance. *Annual Review of Political Science*, *16*(1), 379–398. https://doi.org/10.1146/annurev-polisci-012512-083711

Beauchamp, T. L., & Childress, J. F. (2019). *Principles of Biomedical Ethics* (7th editio). Oxford University Press.

Belgrave, L. L., & Seide, K. (2018). Grounded Theory Methodology: Principles and Practices. In P. Liamputtong (Ed.), *Handbook of Research Methods in Health Social Sciences* (pp. 1–18). Springer, Singapore. https://doi.org/10.1007/978-981-10-2779-6_84-1

Ben-Zeev, D., Schueller, S. M., Begale, M., Duffecy, J., Kane, J. M., & Mohr, D. C. (2015).

Strategies for mHealth Research: Lessons from 3 Mobile Intervention Studies. *Administration and Policy in Mental Health and Mental Health Services Research*, *42*(2), 157–167. https://doi.org/10.1007/s10488-014-0556-2

Benelli, P., & Low, T. (2019). Ethical primary research by humanitarian actors. *Forced Migration Review*, *June*, 28–29. https://www.fmreview.org/sites/fmr/files/FMRdownloads/en/ethics/benelli-low.pdf

Biesenthal, C. (2014). Pragmatism. In David Coghlan & Mary Brydon-Miller (Ed.), *The SAGE Encyclopedia of Action Research* (pp. 648–650). SAGE Publications Ltd. https://doi.org/10.4135/9781446294406

Binet, L., & Saulnier, M. (2019). *Médecins Sans Frontières, evolution of an international movement: associative History 1971-2011*. http://associativehistory.msf.org/sites/default/files/MSF Associative history 1971_2011.pdf

Broussard, G., Rubenstein, L. S., Robinson, C., Maziak, W., Gilbert, S. Z., & DeCamp, M. (2019). Challenges to ethical obligations and humanitarian principles in conflict settings: a systematic review. *Journal of International Humanitarian Action*, *4*(1). https://doi.org/10.1186/s41018-019-0063-x

Cawthorne, D., & Wynsberghe, A. R. Van. (2019). From HealthDrone to FrugalDrone: Value-Sensitive Design of a Blood Sample Transportation Drone. *International Symposium on Technology and Society, Proceedings*, *2019-Novem*. https://doi.org/10.1109/ISTAS48451.2019.8938019

CHS. (2014). *The Core Humanitarian Standard on Quality and Accountability* (First edit). CHS Alliance, Groupe URD and the Sphere Project. http://www.corehumanitarianstandard.org/files/files/Core Humanitarian Standard -

English.pdf

Chun Tie, Y., Birks, M., & Francis, K. (2019). Grounded theory research: A design framework for novice researchers. *SAGE Open Medicine*, *7*, 205031211882292. https://doi.org/10.1177/2050312118822927

Corbin, J., & Strauss, A. (2015). *Basics of Qualitative Research* (4th editio). Sage Publications, Inc.

Cornet, V. P., Daley, C., Bolchini, D., Toscos, T., Mirro, M. J., & Holden, R. J. (2019). Patient-centered Design Grounded in User and Clinical Realities: Towards Valid Digital Health. *Proceedings of the International Symposium on Human Factors and Ergonomics in Health Care*, *8*(1), 100–104. https://doi.org/10.1177/2327857919081023

Cornet, V. P., Toscos, T., Bolchini, D., Ghahari, R. R., Ahmed, R., Daley, C., Mirro, M. J., & Holden, R. J. (2020). Untold stories in user-centered design of mobile health: Practical challenges and strategies learned from the design and evaluation of an app for older adults with heart failure. *JMIR MHealth and UHealth*, *8*(7), 1–14. https://doi.org/10.2196/17703

Creswell, J. W. (2018). *Research Design* (5th ed.). Sage.

Darcy, J. (2004). *Human Rights and Humanitarian Action: A review of the issues*. https://www.odi.org/publications/1558-human-rights-and-humanitarian-action-review-issues

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly: Management Information Systems*, *13*(3), 319–339. https://doi.org/10.2307/249008

de Waal, A. (1997). *Famine Crimes: Politics & the Disaster Relief Industry in Africa*. Indiana University Press.

Eckenwiler, L., Pringle, J., Boulanger, R., & Hunt, M. (2015). Real-time Responsiveness for Ethics Oversight During Disaster Research. *Bioethics*, *29*(9), 653–661. https://doi.org/10.1111/bioe.12193

Evidence Partners. (2022). *DistillerSR*. https://www.evidencepartners.com

Furlong, P., & Marsh, D. (2010). A Skin not a Sweater: Ontology and Epistemology. In D. Marsh & G. Stoker (Eds.), *Theory and methods in political science*. Palgrave Macmillan.

Gordon, S., & Donini, A. (2015). Romancing principles and human rights: Are humanitarian principles salvageable? *International Review of the Red Cross*, *97*(897–898), 77–109. https://doi.org/10.1017/S1816383115000727

Gould, J. D., & Lewis, C. (1983). Designing for usability-key principles and what designers think. *Conference on Human Factors in Computing Systems - Proceedings*, *28*(3), 50–53. https://doi.org/10.1145/800045.801579

Gray, L. M., Wong-Wylie, G., Rempel, G. R., & Cook, K. (2020). Expanding qualitative research interviewing strategies: Zoom video communications. *Qualitative Report*, *25*(5), 1292–1301. https://doi.org/10.46743/2160-3715/2020.4212

Greenwood, F., Howarth, C., Poole, D. E., Raymond, N. A., & Scarnecchia, D. P. (2017). *The Signal Code: A Human Rights Approach to Information During Crisis*. Cambridge, MA: Harvard Humanitarian Initiative. https://signalcodeorg.files.wordpress.com/2017/01/signalcode_final7.pdf

Groupe URD. (2018). *Compas: Qualité & Redevabilité*. Groupe URD. http://www.urd.org/fr/wp-content/uploads/2018/10/Guide_Compas_FR_V2_mail_091018.pdf

Guba, E., & Lincoln, Y. (2005). Paradigmatic Controversies, Contradictions, and Emerging

Confluences. In N. K. Denzin & Y. S. Lincoln (Eds.), *The SAGE Handbook of Qualitative Research* (pp. 191–215). Sage Publications.

Hazard, G. C. (1995). Law, Morals, and Ethics. *Southern Illinois University Law Journal*, *19*(3), 447–458. https://heinonline.org/HOL/Page?handle=hein.journals/siulj19&id=477&collection=journals&index=

Hickman, L. A., Neubert, S., & Reich, K. (Eds.). (2009). *John Dewey between pragmatism and constructivism*. Fordham University Press.

Hilhorst, D., & Schmiemann, N. (2002). Development in Practice Humanitarian principles and organisational culture: Everyday practice in Médecins Sans Frontières-Holland. *Development in Practice*, *12*(3–4), 490–500. https://doi.org/10.1080/0961450220149834

Holden, R. J., & Boustani, M. A. (2021). Design and Usability Methods. In R. J. Holden & R. S. Valdez (Eds.), *The Patient Factor* (1st ed.). CRC Press.

Humanitarian Accountability Partnership. (2010). *The 2010 HAP Standard in Accountability and Quality Management* (Second edition). HAP International. https://reliefweb.int/sites/reliefweb.int/files/resources/260222E062A32EF085257829005BD6FE-Full_report.pdf

Hunt, M. R. (2008). Ethics beyond borders: How health professionals experience ethics in humanitarian assistance and development work. *Developing World Bioethics*, *8*(2), 59–69. https://doi.org/10.1111/j.1471-8847.2006.00153.x

ICRC. (2015). *Ethical Principles of Health Care in Times of Armed Conflict and Other Emergencies*. International Committee of the Red Cross. https://www.icrc.org/en/document/common-ethical-principles-health-care-conflict-and-

other-emergencies

IFRC. (1994). *The Code of Conduct for the International Red Cross and Red Crescent Movement and Non Governmental Organisations (NGOs) in Disaster Relief*. http://www.ifrc.org/Global/Publications/disasters/code-of-conduct/code-english.pdf

IFRC. (2006). *Statutes of the International Red Cross and Red Crescent Movement*. 25th International Conference of the Red Cross. https://www.ifrc.org/PageFiles/88219/statutes-en-a5.pdf

International Organization for Standardization. (2018, March). *Ergonomics of human-system interaction — Part 11: Usability: Definitions and concepts*. ISO 9241-11:2018(En). https://kebs.isolutions.iso.org/obp/ui#iso:std:iso:9241:-11:ed-2:v1:en

International Rescue Committee. (2018). *Obtaining Meaningful Informed Consent*. https://www.principlesinpractice.info/help-library/irc-research-toolkit-obtaining-meaningful-informed-consent

Jackson, S., & Aakhus, M. (2014). Becoming More Reflective about the Role of Design in Communication. *Journal of Applied Communication Research*, *42*(2), 125–134. https://doi.org/10.1080/00909882.2014.882009

KoboToolbox. (2022). *kobotoolbox/kpi*. https://github.com/kobotoolbox/kpi/

Krebs, H. B. (2014). *Responsibility, legitimacy, morality: Chinese humanitarianism in historical perspective*. ODI/HPG. https://cdn.odi.org/media/documents/9139.pdf

Kushniruk, A., & Nøhr, C. (2016). Participatory design, user involvement and health IT evaluation. *Evidence-Based Health Informatics: Promoting Safety and Efficiency through Scientific Methods and Ethical Policy*, *222*, 139–151. https://doi.org/10.3233/978-1-61499-635-4-139

Leaning, J. (1999). Medicine and international humanitarian law. Law provides norms that must guide doctors in war and peace. *BMJ (Clinical Research Ed.)*, *319*(7207), 393–394. http://www.ncbi.nlm.nih.gov/pubmed/10445905

Levac, D., Colquhoun, H., & O'Brien, K. K. (2010). Scoping studies: Advancing the methodology. *Implementation Science*, *5*(1). https://doi.org/10.1186/1748-5908-5-69

Lischer, S. K. (2003). Collateral Damage: Humanitarian Assistance as a Cause of Conflict. *International Security*, *28*(1), 79–109. https://doi.org/10.1162/016228803322427983

Marco, C. A., Lu, D. W., Stettner, E., Sokolove, P. E., Ufberg, J. W., & Noeller, T. P. (2011). Ethics curriculum for emergency medicine graduate medical education. *Journal of Emergency Medicine*, *40*(5), 550–556. https://doi.org/10.1016/j.jemermed.2010.05.076

Martínez, J. C., & Eng, B. (2016). The unintended consequences of emergency food aid: Neutrality, sovereignty and politics in the Syrian civil war, 2012-15. *International Affairs*, *92*(1), 153–173. https://doi.org/10.1111/1468-2346.12502

Martínez, L. L. (2016). Pueblos indígenas y acción humanitaria en América Latina y el Caribe. In I. Mosel, C. Bennett, & H. Krebs (Eds.), *Aproximaciones a la historia del humanitarismo en América Latina y el Caribe*. ODI/HPG. https://cdn.odi.org/media/documents/Aproximaciones_a_la_historia_del_humanitarismo_en_América_Latina_y_el_Caribe.pdf

MAXQDA 2020. (2019). VERBI Software. https://www.maxqda.com

McCaslin, M. L. (2012). Pragmatism. In L. M. Given (Ed.), *The SAGE Encyclopedia of Qualitative Research Methods* (pp. 672–675). SAGE Publications, Inc. https://doi.org/10.4135/9781412963909

Morgan, D. L. (2014). *Integrating Qualitative and Quantitative Methods: A Pragmatic*

Approach. SAGE. https://doi.org/10.4135/9781544304533

Moussa, J. (2014). *Ancient origins, modern actors: Defining Arabic meanings of humanitarianism*. ODI/HPG. https://cdn.odi.org/media/documents/9290.pdf

MSF. (n.d.). *Independent, impartial, neutral*. Retrieved January 4, 2021, from https://www.doctorswithoutborders.org/who-we-are/principles/independence

MSF Ethics Review Board. (2013). *MSF Research Ethics Framework - Guidance Document*. http://hdl.handle.net/10144/305288

Narang, N. (2015). Assisting Uncertainty: How Humanitarian Aid can Inadvertently Prolong Civil War. *International Studies Quarterly*, *59*(1), 184–195. https://doi.org/10.1111/isqu.12151

National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. (1978). *Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research*. Department of Health, Education, and Welfare. http://videocast.nih.gov/pdf/ohrp_belmont_report.pdf

Neubert, S. (2001). *Pragmatism and Constructivism in Contemporary Philosophical Discourse*. http://www.hwf.uni-koeln.de/data/dewey/File/Neubert_Pragmatism_Constructivism.pdf

O'Brien, B. C., Harris, I. B., Beckman, T. J., Reed, D. A., & Cook, D. A. (2014). Standards for reporting qualitative research: A synthesis of recommendations. *Academic Medicine*, *89*(9), 1245–1251. https://doi.org/10.1097/ACM.0000000000000388

OCHA. (2012). *Humanitarian Principles*. UN Office for the Coordination of Humanitarian Affairs. https://www.unocha.org/sites/dms/Documents/OOM-humanitarianprinciples_eng_June12.pdf

Orbinski, J. (1994). Health Impact of Rights Violations in Conf1ict Zones and the Humanitarian

Imperative: Rwanda-A Case Study. *Refuge: Canada's Journal on Refugees*, *14*(5), 5–8. https://doi.org/10.25071/1920-7336.21827

Orbinski, J. (2009). *An Imperfect Offering: Humanitarian Action for the Twenty-First Century*. Walker & Co.

Parsa-Parsi, R. W. (2022). The International Code of Medical Ethics of the World Medical Association. In *JAMA* (Vol. 328, Issue 20, pp. 2018–2021). American Medical Association. https://doi.org/10.1001/jama.2022.19697

People in Aid. (2003). *Code of Good Practice in the management and support of aid personnel* (Second edi). https://reliefweb.int/sites/reliefweb.int/files/resources/DA4AA643ACE54425C1256C6B003B903A-pia-code.pdf

Peters, M. D. J., Godfrey, C. M., Khalil, H., McInerney, P., Parker, D., & Soares, C. B. (2015). Guidance for conducting systematic scoping reviews. *International Journal of Evidence-Based Healthcare*, *13*(3), 141–146. https://doi.org/10.1097/XEB.0000000000000050

Peters, M. D. J., Godfrey, C., Mcinerney, P., Baldini Soares, C., Khalil, H., & Parker, D. (2017). Chapter 11: Scoping reviews. In E. Aromataris & Z. Munn (Eds.), *Joana Briggs Institute Reviewer's Manual* (Issue September, pp. 6–24). https://reviewersmanual.joannabriggs.org/

Pham, P. N., & Vinck, P. (2012). Technology fusion and their implications for conflict early warning systems, public health, and human rights. *Health and Human Rights*, *14*(2), 106–117. http://www.jstor.org/stable/healhumarigh.14.2.106

Pictet, J. (1979). *The fundamental principles of the Red Cross*. http://www.ifrc.org/PageFiles/40669/Pictet Commentary.pdf

Polman, L. (2010). *The Crisis Caravan: What's Wrong with Humanitarian Aid*. Metropolitan

Books.

Raymond, N. A., & Card, B. (2015). *Applying Humanitarian Principles to Current Uses of Information Communication Technologies: Gaps in Doctrine and Challenges to Practice*. Harvard Humanitarian Initiative. https://hhi.harvard.edu/sites/default/files/publications/signal_program_humanitarian_principles_white_paper.pdf

Savoldelli, A., Codagnone, C., & Misuraca, G. (2014). Understanding the e-government paradox: Learning from literature and practice on barriers to adoption. *Government Information Quarterly*, *31*(SUPPL.1), S63–S71. https://doi.org/10.1016/j.giq.2014.01.008

Schopper, D., Dawson, A., Upshur, R., Ahmad, A., Jesani, A., Ravinetto, R., Segelid, M. J., Sheel, S., & Singh, J. (2015). Innovations in research ethics governance in humanitarian settings. *BMC Medical Ethics*, *16*(10), 1–12. https://doi.org/10.1186/s12910-015-0002-3

Schwartz, L., Hunt, M., Sinding, C., Elit, L., Redwood-Campbell, L., Adelson, N., & De Laat, S. (2012). Models for humanitarian health care ethics. *Public Health Ethics*, *5*(1), 81–90. https://doi.org/10.1093/phe/phs005

Simon, H. A. (1996). *The sciences of the artificial* (3rd ed.). MIT Press.

Singer, P. (2021). Ethics. In *Encyclopedia Britannica*.

Slim, H. (2015). *Humanitarian Ethics: A Guide to the Morality of Aid in War and Disaster*. Oxford University Press.

Sphere Association. (2018). *The Sphere Handbook: Humanitarian Charter and Minimum Standards in Humanitarian Response* (4th editio). Practical Action Publishing. https://doi.org/ISBN 978-1-908176-00-4

Sphere Project. (2004). The Sphere Project: Humanitarian Charter and Minimum Standards in

Humaniratian Response. In *Response* (Second edi, Vol. 1). Oxfam Publishing. https://doi.org/10.3362/9781908176202

Stewart, S. (2014). Design Research. In *The SAGE Encyclopedia of Action Research* (pp. 246–247). SAGE. https://doi.org/10.4135/9781446294406

Suzman, J. (2017). *Affluence Without Abundance: The Disappearing World of the Bushmen*. Bloomsbury.

Terry, F. (2002). *Condemned to repeat? : the paradox of humanitarian action*. Cornell University Press.

The Ethics Centre. (2020). *What is Ethics?* The Ethics Centre. https://ethics.org.au/why-were-here/what-is-ethics/?gclid=Cj0KCQiA7aPyBRChARIsAJfWCgJX4UMObXi-8mocGDgBBLPGNH9mc5-cIBBqh-zgFBnX6TBQLpooDoIaAmWlEALw_wcB

Thorley, L., & Henrion, E. (2019). *Ethical standards rapid literature review*. DFID. https://assets.publishing.service.gov.uk/media/5d5e3df5ed915d08d612d835/DFID_Ethics_Literature_Review_FINAL_30_05_19.pdf

Travis, D., & Hodgson, P. (2019). *Think Like a UX Researcher: How to Observe Users, Influence Design, and Shape Business Strategy* (1st ed.). CRC Press.

Trials of war criminals before the Nuernberg military tribunals under Control Council law no. 10 (Volume 2). (1949). U.S. Government Printing Office. https://collections.wip.nlm.nih.gov/catalog/nlm:nlmuid-01130400RX2-mvpart

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D. J., Horsley, T., Weeks, L., Hempel, S., Akl, E. A., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., Wilson, M. G., Garritty, C., … Straus, S. E. (2018). PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. *Annals of*

Internal Medicine, *169*(7), 467. https://doi.org/10.7326/M18-0850

Vaishnavi, V. K., & Kuechler, W. (2015). Chapter 2: Introduction to Design Science Research in Information and Communication Technology. In *Design science research methods and patterns: Innovating information and communication technology* (2nd ed.). CRC Press. https://doi.org/10.1201/9781420059335

Van Mierop, E. S. (2015). Coming clean on neutrality and independence: The need to assess the application of humanitarian principles. *International Review of the Red Cross*, *97*(897–898), 295–318. https://doi.org/10.1017/S181638311500065X

Wade, N. (2006). Before the dawn: recovering the lost history of our ancestors. In *Choice Reviews Online* (Vol. 44, Issue 02). Penguin Books. https://doi.org/10.5860/choice.44-1018

Walker, P. (2005). Cracking the code: The genesis, use and future of the Code of Conduct. *Disasters*, *29*(4), 323–336. https://doi.org/10.1111/j.0361-3666.2005.00295.x

Walker, P., & Purdin, S. (2004). Birthing sphere. *Disasters*, *28*(2), 100–111. https://doi.org/10.1111/j.0361-3666.2004.00246.x

World Medical Association. (2013). World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *JAMA*, *310*(20), 2191–2194. https://doi.org/10.1001/jama.2013.281053

Yeophantong, P. (2014). Understanding Humanitarian Action in East and Southeast Asia: A Historical Perspective. In *HPG Working Paper* (Issue February). ODI/HPG. http://www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/8807.pdf

Chapter 3 Improving Humanitarian Needs Assessments through Natural Language Processing

Tino Kreutzer[1], Patrick Vinck[2], Phuong Pham[2], Aijun An[1], Lora Appel[1], Eric DeLuca[3], Grace Tang[3], Muath Alzghool[1], Kusum Hachhethu[4], Bobi Morris[5], Sandie Walton-Ellery[6], John Crowley[7], James Orbinski[1]

[1] York University

[2] Harvard University

[3] Translators Without Borders

[4] World Food Programme

[5] International Rescue Committee

[6] ACAPS

[7] NetHope

Abstract

An effective response to humanitarian crises relies on detailed information about the needs of the affected population. Current assessment approaches often require interviewers to convert complex, open-ended responses into simplified quantitative data. More nuanced insights require the use of qualitative methods, but proper transcription and manual coding are hard to conduct rapidly and at scale during a crisis. Natural language processing (NLP), a type of artificial intelligence, may provide potentially important new opportunities to capture qualitative data from voice responses and analyze it for relevant content to better inform more effective and rapid humanitarian assistance operational decisions. This paper provides an overview of how NLP can be used to transcribe, translate, and analyze large sets of qualitative responses with a view to improving the quality and effectiveness of humanitarian assistance. We describe the practical and ethical challenges of building on the diffusion of digital data collection platforms and introducing this new technology to the humanitarian context. Finally, we provide an overview of the principles that should be used to anticipate and mitigate risks.

Improving Humanitarian Needs Assessments through Natural Language Processing

The urgent humanitarian needs of people affected by conflict, natural disasters, and climate change have increased significantly in recent years. Today, 69 million people have been forced from their homes, and more than 200 million people need some form of humanitarian assistance (Clarke, 2018, p. 81; UNHCR, 2018, p. 2). Affected individuals and communities are an indispensable source of information about needs, preferences, and existing resources for effective disaster response. Humanitarian organizations predominantly conduct successive quantitative interviews with affected people, both to understand initial needs as well as to monitor, improve, and evaluate the response throughout the program cycle. Today, this is accomplished largely through face to face surveys using mobile data collection applications, while a small but growing number of organizations also use computer-assisted telephone interviews (CATI). Like paper-based methods, these tools require interviewers or respondents themselves to convert complex responses into categorical or numeric data that can then be objectively analyzed with relative ease and speed.

More nuanced insights into the lives of affected people continue to require the use of qualitative survey methods as well as manual coding and analysis that are time-consuming and hard to conduct at scale during any fixed period in an unfolding humanitarian crisis. Humanitarian crisis is understood here as “an event or series of events representing a critical threat to the health, safety, security or wellbeing of a community, usually over a wide area” (WHO, 2007). At the same time, responses to qualitative questions often contain important contextual information not captured in a quantitative survey. For example, food insecurity is common among people impacted by disasters and complex emergencies. Understanding how households cope in such situations is a key objective of many humanitarian needs assessments.

However, an open-ended question such as “how do you cope with the lack of food” often cannot be asked because of the challenges in objectively capturing the response in a manner that can be analyzed to inform operational decision-making. Rather, multiple questions with scale-like response keys are used, which can include a common coping index measure. These, however, may miss important local coping mechanisms and limit respondents’ input, as illustrated in Figure 4. Here, qualitative information was difficult to process, requiring labor intensive and time-consuming manual transcription, translation, and coding. Natural language processing (NLP), a type of artificial intelligence (AI), can provide potentially far-reaching new opportunities to rapidly analyze voice responses for relevant content to inform humanitarian assistance decisions. NLP capabilities are becoming widely available through commercial applications released by companies such as Amazon, Microsoft, Google, and IBM, as well as open source alternatives. However, these AI tools remain largely unavailable in many humanitarian emergency settings where local languages have not (yet) been incorporated by NLP technologies.

Despite rapid technical progress and growing interest, NLP capabilities have not yet been used in humanitarian settings to understand population needs (Emergency Data Science Workshop, 2018). In this paper, we describe a feasibility strategy for use of NLP in humanitarian crises. Our vision is to develop a new way of engaging with affected populations while at the same time providing humanitarian responders with augmented information so that they can respond more effectively and efficiently. As organizations become able to systematically transcribe, translate, and analyze their dialogue with individuals in affected communities—as opposed to merely extracting responses—we suggest that humanitarian assistance may become more effective and efficient, while potentially increasing trust from affected communities. This

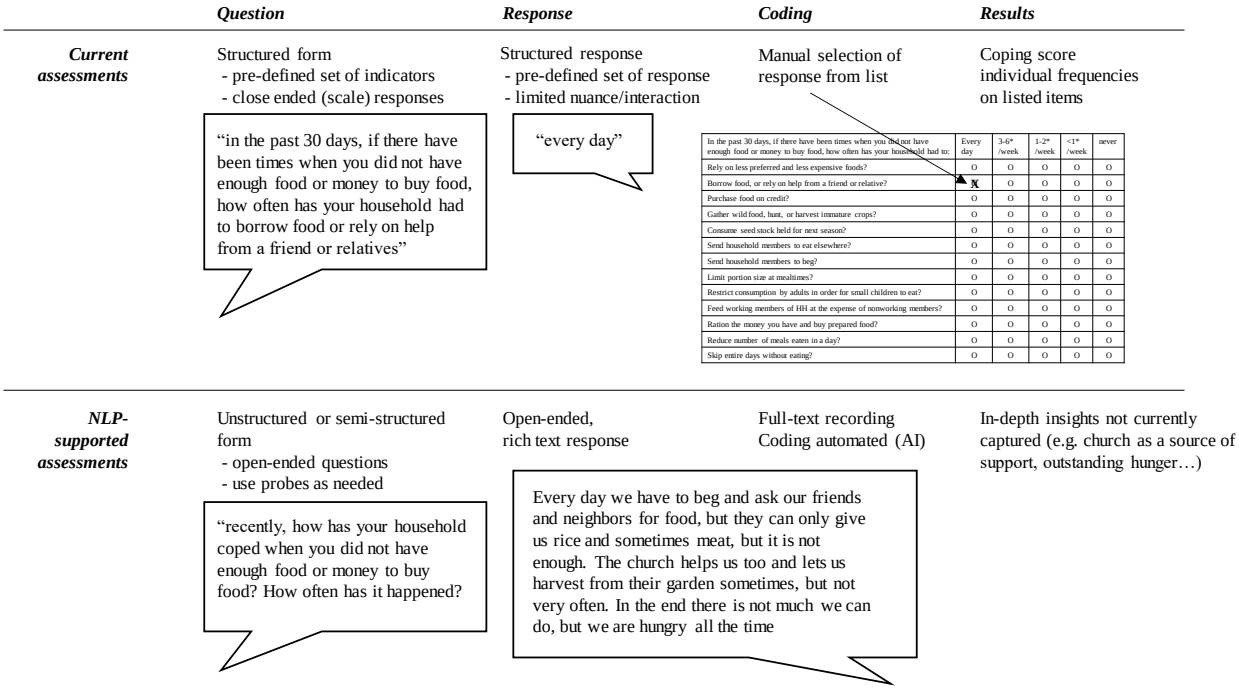
will require identifying the best NLP methods to analyze different kinds of responses, and—for many emergency settings—creating new transcription models for languages not yet amenable to NLP tools.

The number of people in need of humanitarian assistance will likely grow significantly as the effects of climate change intensify and infectious disease outbreaks become harder to control in the growing number of precarious urban settlements (Ghazali et al., 2018). The innovation we describe here may be highly scalable, save a considerable amount of resources, and produce actionable data in close to real time. Already, funding for humanitarian assistance only covers 56% of estimated requirements (OCHA, 2018, p. 8), highlighting the need for innovative strategies and tools to provide rapid, more accurate evidence on how to best use these limited resources.

Employing this new approach is not without ethical challenges. Without upfront and ongoing identification of the socio-political complexity that often leads to or accompanies humanitarian emergencies, and without recognizing the limits and potential biases of NLP techniques, humanitarians may exacerbate context biases that make a particular group vulnerable in first instance, replicate NLP biases, or expose populations to new risks (especially in the domain of security), all with potentially severe consequences for individuals and population groups. Engineers and humanitarian innovators planning to use NLP tools in humanitarian assistance should understand this potentially complex ethical playing field and anticipate and evaluate the potentially harmful consequences that new technologies might bring with them.

Figure 4

Schematic Comparison Between Current Interviewing Methodology and NLP-Supported Assessments



In this paper, we will first describe the limitations of current approaches and technologies being used for primary data collection in humanitarian assistance. Second, we outline our methods and how expected results from using NLP could improve qualitative information in various types of humanitarian data collection. Finally, we examine the need for and implications of the growing array of humanitarian, ethical principles and standards and their implications for the development and use of NLP.

3.1. Understanding People Affected by Crisis Quantitatively

Humanitarian assistance refers to coordinated actions that save lives, alleviate suffering, and maintain human dignity during and after human-made crises and disasters caused by natural

hazards, and is guided by the fundamental principles of humanity, impartiality, neutrality, and independence (Pictet, 1979). Humanitarian assistance here is considered to include “protection”, which “encompasses all activities aimed at obtaining full respect for the rights of the individual in accordance with the letter and the spirit of the relevant bodies of law.” (IASC, 1999) Decisions on necessary assistance and the best methods for delivering it are increasingly driven by a wide array of primary and secondary data to address the needs of the affected population. Humanitarian needs assessments (HNA) and other humanitarian surveys are typically conducted quantitatively during the early stages of a crisis to inform organizations and donors of specific gaps and long-term needs, and during the crisis to monitor and assess progress. This often includes several instances of primary data collection over the first several weeks, either to inform specific sectors (such as shelter needs or access to water and sanitation services) or through a single multi-sectoral initial rapid assessment conducted by multiple organizations (IASC, 2015). These assessments are conducted using structured questionnaires, often with a selection of key informants at the community level. Other types of quantitative data collection, usually occurring after the first two weeks of the crisis onset, may include household-level surveys to assess needs in more detail, to monitor progress of a particular intervention, or to evaluate it after completion. A major advantage of such quantitative surveys is that information provided by the respondent is immediately coded by the interviewer, typically by selecting pre-programmed choices in a questionnaire. Data can therefore be analyzed immediately at the end of data collection—or even meaningfully collated while survey work is still underway.

The growing availability of the Internet, and especially the widespread use of mobile phones, has led to numerous innovations that have changed (and continue to shape) the practice of data collection in humanitarian crises. In recent years, most humanitarian face-to-face surveys

have moved from paper forms to handheld computer-assisted personal interviewing (CAPI) technologies, resulting in high-quality data and faster results (Mock et al., 2013). KoboToolbox, a free and open source platform based on the OpenDataKit (Anokwa et al., 2009), was embraced in 2014 as the preferred humanitarian survey tool by the UN Office for the Coordination of Humanitarian Affairs (OCHA), leading to widespread adoption by a broad range of international and national humanitarian agencies (OCHA, 2015). As of 2019, public KoboToolbox servers received more than 73 million survey submissions and had more than 200,000 users (figures from the authors). For example, 63% of humanitarian organizations working in Syria in 2017 reported using KoboToolbox for primary data collection (Building Markets & Orange Door Research, 2018, p. 19). Similar to traditional paper-based surveys, these tools are generally employed in face-to-face interactions, requiring staff to travel to sampled locations. Under ideal circumstances, a single interviewer can conduct, for example, ten 30-minute household-level interviews per day in a dense urban environment. However, humanitarian crises often pose extreme limits on interviewers' ability to travel, either for logistical or security reasons. This can significantly slow down face-to-face data collection in practice. A sharp increase in violence against aid workers (Humanitarian Outcomes, 2021) in recent years has further hampered physical access to affected areas. For some contexts, such as in the Pacific region, access to certain communities requires long and expensive travel even in non-disaster circumstances. Time, security, and budgetary limitations sometimes force humanitarian organizations to exclude some groups and communities from face-to-face surveys, leading to large information gaps and unrepresentative data.

In an effort to overcome these challenges, a number of surveys in low- and middle-income countries have moved from face-to-face interviews to phone-based interviews (Gibson et

al., 2017). Cell phone networks can now be accessed by 96% of the world's population (ITU, 2018, p. 8), and smartphones are seen as an essential lifeline among displaced groups (Poole et al., 2017; UNHCR, 2016). Remote data collection most commonly involves computer-assisted telephone interviews (CATI), which are conducted by a trained interviewer (often a call center operator) who asks questions in the local language using structured questionnaires. Similar to face-to-face interactions, responses are classified immediately in close-ended categories to facilitate rapid quantitative analysis. The World Food Programme (WFP) has conducted CATI surveys in 31 humanitarian crises to date to collect critical food security and nutrition information (Morrow et al., 2016; Robinson & Obrecht, 2016). Other forms of remote phone-based interviews are conducted by text messages and Interactive Voice Response (automated either through voice recognition or responding by pressing a number on the phone keypad)—all of which are used primarily for collecting quantitative data (Lamanna et al., 2019, pp. 2–3).

The key advantages of CATI surveys are speed and precise categorical data to inform immediate humanitarian operational decision making. Compared to face-to-face interviews, CATI methods are also more practical for collecting information in physically inaccessible areas and less costly for conducting real-time monitoring (Bauer et al., 2013). For example, WFP estimates $3-9 per complete CATI survey vs. $20-40 for a complete face-to-face survey. The most obvious limitation of CATI methods is their inability to reach people who may not have access to mobile technologies (either directly or through a family member). This particularly includes rural populations as well as women, elderly, and disabled people as the only mobile phone is frequently controlled by the male household head. In low-income countries, cell phones are often shared across a larger network of friends or family members (Kreutzer, 2009), while ownership is skewed towards males and urban households with higher income (Leo et al., 2015,

p. 14). This challenge can be addressed by using complex statistical weighting procedures to reduce bias in the data, which requires reliable demographic baseline data and a minimum level of access to cell phones among the population (Leo et al., 2015). Information provided by respondents may also depend on the type of data collection method used, whereby face-to-face interviews can be both conducive and a hindrance to more honest responses, depending on the topic (Dette et al., 2016, p. 23; Lamanna et al., 2019; Langhaug et al., 2011). A comparison between face-to-face and remote data collection methods is shown in Table 6.

Table 6

Comparison Between Face-to-Face and Remote Interviewing Methods

	Face-to-Face Interviews	**Remote Interviews**
Technology	CAPI (e.g., KoboToolbox, OpenDataKit, CommCare) for quantitative; digital audio recording or personal notes for qualitative methods	CATI call centers; Interactive Voice Response, text messages
Example use cases	Humanitarian needs assessments, in-depth, or cross-sectional surveys	Rapid surveys with fewer questions (needs assessments, situation monitoring, program monitoring)
Advantages	Allows longer interviews; representativeness; no bias due to technology access	Faster data collection/shorter turnout time; allows more frequent data collection and larger samples; cheaper; allows the collection of data from hard to access and insecure areas
Disadvantages	Expensive, slow, restricted by physical access	Exclusion of people without access to mobile phones or mobile networks; exclusion of people with low literacy for text message surveys; not suited for long and complex surveys
Qualitative data	Can be captured, but rarely done with proper transcription, translation	Can be captured, but rarely done with proper transcription, translation

3.2. Understanding People Affected by Crisis Qualitatively

The key limitation of both face-to-face and CATI surveys in humanitarian emergencies is the lack resources to properly handle qualitative information. There have been many initiatives over the past three decades to engage more meaningfully with the people affected by conflicts and disasters by complementing quantitative surveys with qualitative data collection methods. These initiatives include focus group discussions with affected populations to evaluate the effectiveness of a specific program or using hotlines to collect feedback from program participants. But the use of these methods to better inform humanitarian assistance or give affected people a greater sense of ownership still varies widely across different emergencies (HHI and ICRC, 2018, p. 12,40). In these instances, the information provided by respondents or participants is typically captured through handwritten notes, and—in the case of focus group discussions—through audio recordings for later analysis. Nuanced human expression and thinking cannot be easily captured by quantitative methods. An important reason for the limited use of qualitative methods is the time and cost associated with transcription, translation, and content analysis. For example, ten unstructured interviews of 30 minutes each may take 50-100 hours to transcribe and translate, while analysis—even with the help of software—can require an additional 20-40 hours.

Both CATI and CAPI can be used for qualitative methods, such as for recording complex open-ended responses in a largely quantitative survey or even entire unstructured or semi-structured interviews. However, most surveys that inform humanitarian operations, monitor progress, or track opinions about aid actors, are conducted using quantitative research instruments. Even in instances where qualitative information is collected, it is rarely systematically analyzed. The main reasons for this limitation are the time and cost associated

with using qualitative data well: Transcribing, translating, and coding open-ended responses should ideally be done by trained professionals who process collected data in the language of the affected population. Already, there are too few staff with such skills. Fees for trained translators, transcribers, as well as staff with experience in qualitative analysis can quickly exhaust small budgets. Likewise, the time needed for these activities is often a multiple of the original interview response, creating a lag of days—and often weeks—between data collection and final analysis. Instead, organizations are forced to use a single staff member with limited training to cover all these tasks—or collect fewer qualitative data to begin with.

Table 7

Comparison Between Existing Quantitative and Qualitative Data Collection Methods

	Quantitative	**Qualitative**
Example of humanitarian data collection	Household survey on nutrition needs	Semi-structured impact evaluation interviews; feedback collection hotlines
Technologies for data collection	Paper, CAPI, CATI	Paper based note taking, digital audio recording
Speed for data to be available	Fast: Can be coded immediately by the interviewer	Very slow and expensive to scale: Requires recording, transcription, translation, manual coding, and post-hoc analysis
Information depth	Complex questions are hard to code or summarize during the interview (and attempts by interviewers can be very unreliable)	Allows for deeper analysis of knowledge, sentiments, perceptions

These challenges, summarized in Table 7, prevent organizations from fully using qualitative methods at scale, particularly at the early stages of emergencies when representative survey samples require a large number of interviews and when rapid analysis is essential. Complex questions requiring careful qualitative analysis are therefore largely lacking in humanitarian surveys. Instead, a common approach is to replace them with simpler, less nuanced

alternatives using closed-ended categorical or ordinal response options. Interviewers are required to interpret and categorize the information immediately, turning each response into a single variable. But this approach also bears significant risks: Interviewers are often hired rapidly in a crisis and given only minimal training. Similar responses may be coded differently and incorrectly, depending on the interviewer's biases and comprehension of key terms used.

When considering survey methods, significant linguistic challenges can emerge in many of the most urgent humanitarian crises. Human languages can vary in dialect, morphology, grammar, syntax, and semantic structure—all of which define and affect meaning which itself can be culturally specific, and which can change over space and time (Deutscher, 2010). Linguistic forms and meanings can evolve with use in a given culture, and as cultures and their languages interact. Ensuring that the original meaning is fully captured in the target language can be extremely challenging, particularly for qualitative information (Al-Amer et al., 2016; Bowden & Fox-Rushby, 2003). These challenges often make translation from a source to a target language an ongoing interpretive process—even for dedicated professionals.

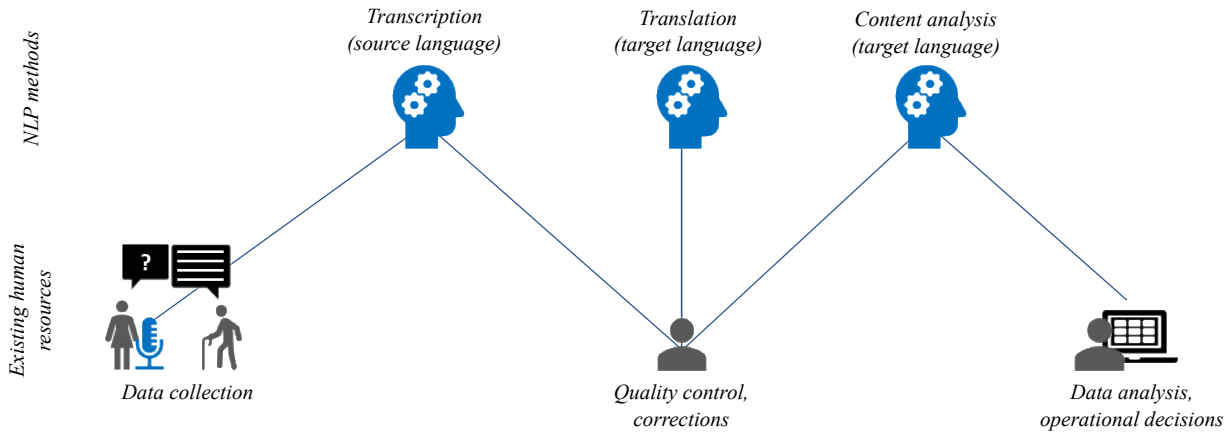
There are over 3,000 languages spoken in the 42 countries currently experiencing humanitarian crises or situations of concern (Translators Without Borders, 2019). For example, there are over 40 languages spoken in the six conflict-affected states of northeast Nigeria alone (Translators Without Borders, 2018b). National citizens hired as humanitarian interviewers in Nigeria are often asked to read surveys in English, sight translate them into Hausa or a variety of other local languages, and then instantly try to match the respondent's answers with one of the listed categories in English. One study with humanitarian interviewers found that these challenges are pervasive even among highly experienced staff (Translators Without Borders, 2018a, p. 6). Further, a recent review found that two quantitative assessments in the Rohingya

crisis yielded wildly different results on the same indicators (ACAPS, 2019). Such findings lead to the question of how the quest for speed may have rendered many humanitarian assessments less reliable.

Even in contexts where there is no significant language diversity, there can be real challenges. Nearly all of the refugees in Cox's Bazar, Bangladesh, speak Rohingya, a language very closely related to Chittagonian, which is spoken by nationally hired humanitarian field staff. Yet, a study found that the differences between the two languages are significant enough that nearly a third of Rohingya refugees were unable to understand a basic sentence in Chittagonian (Maya Hasan, 2018). And since both Rohingya and Chittagonian lack formalized written scripts, survey instruments are often written in two or more languages, sometimes even offering transliterations of Rohingya using Bengali script to help local responders pronounce the questions appropriately. These workarounds are not simple and require significant training and support to maintain an often-elusive consistency.

Figure 5

Schematic Overview of NLP-Supported Humanitarian Assessments



There has been some “bottom-up” innovative data collection in humanitarian assistance. These include crowdsourcing based on text messages (Meier, 2014), feedback collection on aid deliveries (Bonino et al., 2014), voluntary reporting on conflict events (Van der Windt & Humphreys, 2016), and social media used to communicate personal needs (Facebook, 2017). The growing number of humanitarian call centers, which allow affected people to find reliable information or communicate specific complaints with aid delivery, are also included in this category (such as during the 2014–2016 Ebola outbreak in Sierra Leone or for displaced people in Iraq (Dette et al., 2016, p. 21)). However, for survey purposes these methods generally do not allow for randomness in the sample, are only practical for a small number of questions, and require large numbers of staff to transcribe and categorize unstructured audio or text data. In short, none of the innovations to date enable a reliable, rapid and effective use of qualitative data, overcoming the challenges identified above. NLP may offer an opportunity to address this gap.

3.3. Using NLP to analyze qualitative data in humanitarian assistance

Artificial intelligence (AI) can be defined as scientific and technical attempts to build machines that act rationally, with the capacity to mimic human cognitive functions to perceive, understand, predict, or manipulate (Russell & Norvig, 2009, pp. 1–30). When such techniques are deployed to real-life contexts, they are often referred to as AI systems (AIS). AIS are understood as “software (and possibly also hardware) systems designed by humans that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data and deciding the best action(s) to take to achieve the given goal” (High-Level Expert Group on Artificial Intelligence, 2019a, p. 6). Because of the vagueness of many definitions of AI, the Institute of Electrical and Electronics

Engineers (IEEE) proposed using the more narrow term “autonomous and intelligent systems” (A/IS). (The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems, 2019)

In recent years, Natural Language Processing (NLP) has received increasing research and commercial attention. NLP is understood as “the use of computational methods to analyze and process spoken or written statements in a language commonly used by humans” (Assal et al., 2011, p. 2). Breakthroughs in transcribing, translating, and understanding human speech have been propelled by innovations in the field of AI, including data mining, machine learning, deep learning, and reinforcement learning (Sajad Mousavi et al., 2018). The complex tasks of machine translation and natural language understanding have, in turn, been based on advances in information extraction and related subfields (Singh, 2018). The rapid advances in these fields have led to a boom in commercial applications for digital assistants such as Apple’s Siri, Google Assistant, and Amazon Alexa. At the same time, open source software and training models are being created to replicate or surpass these systems’ performance.

The effective use of NLP may enable humanitarian assistance organizations to vastly increase the use of qualitative methods in their interactions with affected people to better enable affected people to communicate their needs. Doing this requires linguistic, technological and methodological approaches for three separate fields: (1) transcription of voice data into written language, (2) translation from a source language into a target language , and (3) various types of NLP analysis of target language transcripts of what was said, as displayed in Figure 5.

3.3.1. Transcription

Proper transcription from audio recordings can be challenging even for skilled annotators. Today, transcription software is already widely available through commercial services such as from Google, IBM Watson, Microsoft Azure, Amazon Web Services, and open source

alternatives such as Transformer (Vaswani et al., 2017), DeepSpeech (Hannun et al., 2014), and OpenNMT (Klein et al., 2017). For example, Google's speech-to-text service supports 64 languages (Google Cloud, n.d.). The platform lists 120 "languages and variants", but 54 refer to dialects of the same language (e.g., Australian English vs. American English).

The most technical challenge for automating transcription (also referred to as speech recognition systems) is due to the informal nature of survey responses. Spontaneous speech is an "unplanned, non-rehearsed, naturally occurring, and non-experimental type of speech that forms the means of communicating information between individuals" (Alzghool, 2009, p. 1). State-of-the-art speech recognition technologies have achieved high recognition accuracy for read texts or constrained-spoken interactions (such as broadcast news). However, accuracy is still rather poor for spontaneous speech, which is often not well structured and contains many disfluencies, leading to a higher error rate for automatic speech recognition systems and to redundant information. The four most popular disfluencies are: filler words, repetitions, repairs, and restarts (Furui et al., n.d.). Repetitions are redundant pieces of information that occur when the speaker pauses for a while, considering what to say next, and then repeats the previous information. Repairs occur when the speaker says something wrong and corrects themselves immediately. Restarts occur when a whole part of a sentence is abandoned, and the speaker starts another one. Personal speaker characteristics also affect transcribed text, such as heavy accents, age-related co-articulations, speaker and language switching, and emotional speech.

For this reason, human intervention to correct falsely transcribed speech remains necessary to ensure that the transcript is accurate—even under ideal circumstances (such as formal speech, high audio recording quality, or highly advanced NLP language models).

3.3.2. Translation

As described above, correct translation of qualitative information through human interpreters faces numerous challenges. Advances in deep learning over the last decade have led to a rapid increase in the quality and quantity of available machine translation tools. For example, the quality of machine translation systems for commercially viable languages such as Chinese is now nearing parity with professional human translators in particular domains (Microsoft, 2018). Typically, machine translation first requires creating a written transcript in the source language which is then translated into a target language. However, recent methods also suggest that it may soon become feasible to directly translate source audio into the target language text (Bansal et al., 2018; Weiss et al., 2017).

Similar to the limitations of automatic speech recognition systems, trained translators are still required to correct NLP-generated translations, ranging from obvious terminology mistakes maintaining semantic nuances existing in the source language. For historical and commercial reasons, however, there are very few languages in countries affected by humanitarian crises for which machine transcription and translation is available, such as Hausa, Rohingya, Swahili, Fulfulde, or Amharic (Abbott & Martinus, 2018; Gu et al., 2018). Despite a recent willingness from companies such as Google, Microsoft, and Facebook to also support these languages, there remains insufficient training data for machine learning in many of these languages to satisfy the more data-intensive approaches of neural machine translation (Koehn & Knowles, 2017).

These languages also often have little to no existing text or audio datasets that can be used for NLP training purposes. Open source NLP software, coupled with crowdsourcing of labeled recordings, are becoming a promising option to bridge this new digital divide by creating open source models for source languages common in contemporary humanitarian emergencies.

For example, the Common Voice platform created by Mozilla makes it possible to conduct speech recognition in many of the world's languages, with volunteers already having contributed 2,198 hours of training data across 29 languages as of June 2019 (CommonVoice, 2019). Efforts to collect more training data for three priority humanitarian languages are also underway among several humanitarian partners, led by the nonprofit organization Translators Without Borders (Ansari & Petras, 2018). However, the lack of training data in these source languages is not a quick problem to solve. At a minimum, it involves a coordinated effort to engage with communities of linguists to actively build and improve these datasets. This can be a challenging task for languages such as Rohingya that lack a formalized translation industry or experienced translators.

3.3.3. Analysis

For our purposes, qualitative text analysis refers to the process of establishing the content and meaning of interview responses and other unstructured data. This has traditionally been done by humans at a linear scale (more text requires correspondingly more time), and only with limited use of software, such as NVivo. However, recent advances in multiple NLP disciplines have led to an explosion of the tools available to analyze large amounts of data, including commercial platforms such as Google's Cloud Machine Language Engine or IBM's Watson Natural Language Classifier. A growing number of sophisticated open-source NLP tools can also be used to process qualitative survey data, depending on the type of questions asked and the data sought from responses. Examples include the Natural Language Toolkit for classification (such as the severity of needs) (Loper & Bird, 2002), TextRank for summarization (such as identifying most urgent needs) (Mihalcea, 2004), GATE for information extraction (such as the respondent's age) (Cunningham et al., 2013), and MALLET for topic modeling (such as understanding a

household's top priorities for disaster recovery) (McCallum, 2002). Some of these methods rely on supervised machine learning, which requires creating a human-labeled dataset of real-life responses (e.g., education level) that results in a model for predicting how future responses can be classified. Unsupervised learning, such as for clustering responses based on similar content, does not require data specifically labeled by humans, but instead solely relies on the data collected. Each approach needs to be carefully calibrated to overcome a long list of challenges, such as the use of double meaning or corrections to what was said previously.

There remain several challenges and limitations to employing this approach for surveys: First, the accuracy of NLP analysis varies significantly depending on the amount of data available for training new models and the amount of variation in survey responses. Human verification of any NLP-generated data is thus essential to correcting mistakes and improving the model's algorithms. Second, many existing NLP analysis tools require English text as input. As a result, responses collected in other source languages usually need to be translated into English as a target language, adding an additional potential for errors. Alternatively, it is possible to train some NLP toolkits to work in other source languages, but this approach requires substantial resources to generate high-quality models. In such cases, analysis can be done in the source language first, thereby only translating coded results into English (or other languages), if necessary.

3.3.4. Process

This section outlines a proposed functional design of NLP-supported humanitarian needs assessments (as well as similar types of primary data collection) and describes some of the changes needed to adapt existing processes. First, during initial data collection, audio recordings for each question/response pair would need be saved on a mobile device (for face-to-face

interaction in CAPI) or on a call center computer (for CATI), along with proper timestamps to mark the beginning and end of each response. The recordings would then be transmitted to a server for further processing.

Second, audio recordings in the source language would be converted into text using an NLP transcription model. Third, in many cases, the transcribed text in the source language would then be translated into a different target language (such as English) for analysis. Quality-control by native speakers and professional translators should be ensured to correct semantic and terminological mistakes stemming from the automatic transcription and translation steps.

Fourth, content analysis based on transcribed, translated, and human-corrected text would require a combination of different NLP analysis techniques (as mentioned in the previous section). Ideally, unsupervised NLP methods should be used wherever possible to reduce the need for both training data and creating custom analysis models. Where training data is required for supervised machine learning methods, enumerators would need to categorize (or label) a sufficient set of initial responses, either at the time of the interview or by reviewing the audio recordings at a later stage. As a fifth and final step, trained specialists should conduct quality-control of all NLP-generated categorizations to correct mistakes and improve machine learning models in the process.

The result of this proposed functional design would be a systematically coded dataset that extracts, classifies, and clusters information from spontaneous speech in response structured and unstructured data collection methods. More research is required to 1) test the performance of various NLP toolkits for the same qualitative analysis task; 2) establish standard approaches for analyzing common question types (including by combining different NLP tasks); and 3) establish

the feasibility of creating human-labeled training datasets during ongoing humanitarian assistance operations.

3.3.5. Anticipated Benefits

We expect that creating a system to utilize NLP in humanitarian emergency settings would significantly improve the quality of information collected for humanitarian operational purposes. First, by integrating more open-ended questions into surveys or using more qualitative methods overall, humanitarian organizations would be able to gain a more nuanced and accurate understanding of the topic under investigation. Second, as shown in Figure 4, questionnaires can be designed to include fewer questions, making the interview less rigid and flow more naturally. Instead of asking questions that each result in a single variable needed for analysis, questions can be asked in more open-ended ways and followed up with probing questions as needed. This would invite respondents to elaborate on a topic, giving them more of an opportunity to describe what is important to them. Later analysis of these responses is likely to contain much more information (which can be coded into subsequent survey variables) than what is possible with the current question-answer style of quantitative interviews. Third, standard surveys—even those that are primarily quantitative—could eventually be conducted more quickly, as interviewers would not need to spend time entering responses into pre-designed multiple-choice options or text boxes. Instead, they can simply focus on conducting the interview and move on to the next question as soon as a satisfactory response has been received, as all audio is recorded for later processing.

Shorter interviews are primarily in the interest of crisis-affected respondents, who can return to their important post-disaster activities more quickly. However, they also benefit humanitarian organizations who can increase the number of sampled respondents (thereby

increasing the representativeness for smaller sub-groups or regions). Time savings are much greater, of course, for qualitative methods where manual transcription and content analysis would otherwise require hundreds of hours of staff time. Better information and fewer resources used to generate it should result in a more effective, efficient and nuanced response to needs and a reduction in suffering of people impacted by crises.

The proposed system would substantially augment and integrate well with existing techniques and technologies for field data collection used in humanitarian assistance. As described earlier, the vast majority of humanitarian data collection is conducted between a human interviewer and an individual respondent, either face-to-face using CAPI or remotely in CATI surveys. CAPI technologies, such as KoboToolbox or the OpenDataKit, so far only allow recording audio responses to individual questions, though work is underway to automatically record background audio without the need for interviewer interaction. Similarly, many organizations using CATI already routinely record calls for quality assurance. Call centers lend themselves in particular to this planned system, as limitations in the field on uploading potentially large numbers of audio files through slow Internet connections are avoided. However, as the next section will show, we also need to focus on the potential risks of using NLP in the context of humanitarian crises.

3.4. Anticipating and Mitigating Ethical Challenges

NLP and AIS in general can pose significant new risks to people affected by humanitarian crises. This risk stems both from the increased use of information and communication technologies (ICT) to collect and process ever more detailed personal data, but also from entirely new technologies that can automate analysis and decision-making. Increasing availability of mobile technology, access to the Internet, and the use of social media have led to

considerable enthusiasm to deploy new ICT as an answer to many complex societal challenges, including for humanitarian assistance (Harvard Humanitarian Initiative, 2011; Meier, 2011, 2014). Growing interest by technology companies in the field of disaster response has only heightened this trend, including forays by Google and Facebook that aim to transform humanitarian operations. Over the last several years, there has been a growing awareness around the ethical risks of increased use of ICT in humanitarian crises, as well as calls to mitigate them (Pham & Vinck, 2012; Sandvik et al., 2014; Scott-Smith, 2016). Many specific risks have been identified—both from successful and failed technology deployments (Jacobsen, 2015). Among many issues, this includes remote survey respondents who may be targeted by armed groups (Building Markets & Orange Door Research, 2018, p. 19), wrong decisions based on an erroneous sense of accuracy (Hunt et al., 2016), misleading information sourced from social media and crowdsourcing (Crawford & Finn, 2015), increased vulnerability by excluding those who do not own phones from participation (Poole et al., 2017) and by using exploitive data mining practices (Greenough et al., 2009), and worsening power imbalances between responders and affected people (Sandvik & Raymond, 2017). One particular area of concern has been the premature introduction of novel technologies during humanitarian crises, using crisis settings as a testing ground for experimentation (Sandvik et al., 2017). Humanitarian practitioners have created a number of specific technical guidelines in recent years to address such concerns. A notable example among these is the ICRC *Handbook on Data Protection in Humanitarian Action*, which is the most detailed and stringent resource to date (International Committee of the Red Cross, 2017). Nonetheless, careful balancing of risks and benefits of particular ICT options (including the decision to not use any technology) remains rare (Dette, 2018).

The use of remote data collection technologies faces additional risks, especially when used to enabled remote project management in conflict environments. Here, ICT for remote surveys and other forms of information transmission (including apps such as WhatsApp or Signal) are increasingly used in violent conflicts such as in Syria where many aid organizations were not allowed to operate openly, enabling humanitarian actors to engage with affected people from afar (Steets et al., 2016). While justified by the need to keep (international) staff safe in conflict environments, remote management can also be motivated by cost reduction or convenience reasons, which clashes with the valued principle that proximity to populations in need is essential for humanitarian action.

Data collection technologies supported by AI face yet another layer of risks. Applying such systems to understand people and their behavior has been shown to reinforce or even exacerbate human biases: If the AI training data reflects existing social biases or is constructed based on easily accessible but unrepresentative data, the results can not only repeat but amplify social inequities. For example, commercially available face recognition software performed very poorly for dark-skinned women while excelling with pictures of white men (Buolamwini, 2018), while voice recognition was found to be less accurate for women and speakers of minority dialects (Tatman, 2017). Many of these often-proprietary tools do not publish their source code and underlying training data. As a result, many well-intentioned systems have resulted in negative or controversial results due to different forms of biases that were only discovered after they were deployed (Molnar & Gill, 2018; Paul et al., 2018). However, aiming for greater transparency can also pose risks: The authors of a recently created system that can write complete news articles given only a headline decided not to publish their source code, fearing it might be used to spread “fake news” maliciously (Vincent, 2019).

NLP used to transcribe and analyze human speech is associated with very low immediate risk, but if such information is used to make recommendations for human operators—as in the case of German immigration officials using NLP to detect identity fraud among refugee claimants (Wood, 2018)—the level of potential risk increases significantly. In general, for example, a model could perform better for young, educated, urban males (whose voice or written data may be more easily accessible), thereby leaving women and many other members of society misrepresented in surveys. In order to counter the risk of entrenching societal biases in NLP methods, it is essential to include a large demographically representative training sample used in the process of creating new models (especially women, the elderly, and minorities)—and to use data weighting that increase the use of outliers in order to optimize the performance for all likely users (Amini et al., 2019). Such a process can only reduce biases, not eliminate them. It is therefore important to rigorously document the methods for establishing training data, such as publishing the demographic composition of speakers used for compiling voice training data and to continually test model outputs to identify and combat bias. Imbalances in carelessly collected AI training data can lead to unintended biases that can have severe real-life consequences, making it imperative to anticipate, measure and mitigate these early on.

A recent review has found that existing ethical principles may have little impact in reality as software developers were found to ignore ethics codes in a behavioral experiment (McNamara et al., 2018). Recent scandals involving Facebook and other companies underline the challenge of embedding and enforcing ethical frameworks and practices in everyday business decisions, especially where unintended consequences may not be apparent for months or years.

Scandals involving AIS and controversially acquired personal data received significant attention in 2018 (AI Now Institute, 2018). Such concerns about the ethical implications of AI

tools have only recently been recognized as research priorities. Public concerns around the implications of AI (both current but especially in the future) (Brundage et al., 2018) have resulted in many public policy initiatives to address a wide array of concerns. National governments have largely focused on making AI development a national priority for economic competitiveness, framing it as a national security issue and providing increased research funding. For most countries, there are as yet no laws or policies on the ethical development and use of AI (Future of Life Institute, n.d.). The European Commission has set up an independent panel to establish specific policy recommendations to this end. In April 2019 the panel released a set of ethical guidelines and operational recommendations for creating "trustworthy AI" with the goal of creating concrete policy recommendations (High-Level Expert Group on Artificial Intelligence, 2019b). In May 2019, 42 countries (including all 36 members of the Organisation for Economic Co-operation and Development) adopted an intergovernmental policy guideline to ensure AI systems are designed "in pursuit of beneficial outcomes for people and the planet" (Recommendation of the Council on Artificial Intelligence, 2019, para. 1.1).

Other public policy initiatives include the 23 Asilomar AI Principles as well as the *Montreal Declaration for Responsible Development of Artificial Intelligence* which hopes to "be translated into political language and interpreted in legal fashion" (Montreal declaration, 2018, p. 10). There are also several examples of private sector initiatives to guide managers and engineers to remain ethical. Some of these are brief, non-specific and at a high level, such as Google's *AI Principles*, and the industry-civil society collaboration Partnership on AI's *8 Tenets*. Others provide more detailed guidance, particularly the Association for Computing Machinery, Institute of Electrical and Electronics Engineers (IEEE), and Microsoft. Separately, the humanitarian sector has developed its own array of principles and operational guidance around data and

technology in the field. However, practical implementation remains challenging (Hilhorst & Schmiemann, 2002; Raymond & Card, 2015) while none have addressed AI to date.

It is essential for the humanitarian sector to interact with the growing body of ethical AI principles as well as building on the work done to improve data protection with "traditional" ICT. Private and public sector initiatives to provide guidance for ethical development and deployment of autonomous and intelligent systems should where possible, be used by humanitarian innovators and data scientists instead of creating new guidelines. The use of ethical review boards is essential for shaping, ensuring and enforcing responsible behavior internally. The increasingly standardized and professionalized governance structures can serve as platforms to promote and review the responsible use of AIS. These include for example, the Inter-Agency Standing Committee, NGO associations such as InterAction or the International Council of Voluntary Agencies, as well as the cluster coordination system under the United Nations Office for the Coordination of Humanitarian Affairs.

3.5. Conclusion

Current state-of-the-art humanitarian assessments do not capture the complex and rich context in which humanitarian crises unfold, particularly the experience, needs, and resources of affected communities. Furthermore, they rely on a primarily extractive quantitatively oriented methods that can miss nuanced qualitative information. Routinely analyzing rich interview and dialogue data could help generate more tailored assessments, improving disaster response, and would arguably improve the relationship between communities and the agencies that seek to assist them. NLP methods offer unique capabilities to systematically transcribe, translate, and analyze interview responses. Advancing the use of NLP, however, will first depend on research to establish the viability of the five-step proposed functional design of an NLP tool, described

earlier. It will also require the support of many partners, including UN agencies, international and local NGOs, donors, as well as private sector organizations. It will also require ensuring the ethical design and use of the NLP tool in a manner consistent with the specific complexities and needs of humanitarian crises. Table 3 briefly describes the components of a proposed pilot phase in a humanitarian crisis for which no transcription and translation NLP model (commercial or open source) exists so far.

NLP should complement, not replace face-to-face interviews: Surveys conducted in person allow much more control over the interviewing environment, are able to extend over a longer time (e.g., for cross-sectional studies), and can go into more depth than phone interviews. Further, in-person interviews offer a more personable way to interact with people who have suffered trauma and are struggling to recover. Increasing the usage of existing NLP methods in humanitarian assistance operations has enormous potential benefits by enabling better two-way communication through which affected people can better communicate their needs. We hope that the proposed approach allows humanitarian responders (as well as other survey-intensive domains, such as public health) to rethink the current methodological paradigm that holds that qualitative methods are not compatible with large population samples.

There is an urgent need to bridge the growing gulf between the people affected by humanitarian emergencies and response professionals through improving the quality and quantity of information provided by the affected population. Research and software development are needed to make NLP technology relevant, accessible (and free) for all humanitarian assistance organizations in all crises globally. However, new transcription, translation, and analysis models may not yet prove sufficiently accurate for large-scale deployment in the complex environments that often characterize humanitarian emergencies. Regardless of success, lessons learned on bias

reduction and automated classification, as well as the practical results such as trained transcription models, will be immensely useful for humanitarian assistance and academic research, as well as the overall use of NLP in applied domains.

Table 3

Components of a Proposed Pilot Phase in a Humanitarian Crisis for Which No Transcription and Translation Models Exist

Workstream	Activities
Preparation	Together with relevant partners, select the assessment questionnaires most appropriate to use
	Modify assessments to include more open-ended questions to cover more information needs
Generate transcription and translation model	Manually transcribe and translate sample responses (related to humanitarian assistance and general domain language) to serve as training data
	Collect speech recordings from volunteers as audio training data (from volunteers recruited on the ground and online). Voice collection will specifically seek out women, older age groups, speakers of minority dialects, among others to avoid biases favoring urban young men.
	Verify accuracy and correct as needed, using crowdsourcing and professional translators
	Use weighting methods to train the transcription model as accurately as possible for all population groups
Create analysis models	Transcribe samples of interview responses (first manually, later using transcription model), using categorization from two separate reviewers as labels
	Measure the accuracy of interviewer classification through secondary qualitative analysis to establish the human accuracy rate.
	Use and compare different NLP approaches to achieve (or surpass) human accuracy rate
Toolkit	Create a methodological toolkit that can be applied to all humanitarian data collection contexts and integrated with other methods (CATI, mobile data collection)
	Provide detailed recommendations for replication and scaling up of approach in other emergencies
	Release all documentation, language models, classification algorithms, and software code through open source licenses
Supporting research and documentation	Identify practical, data governance, and ethical issues that need to be addressed by future humanitarian AIS using NLP
	Conduct a scoping review of prevailing NLP algorithms and training data
	Test available transcription models to document existing biases
	Evaluate and publish all results

3.6. References

Abbott, J. Z., & Martinus, L. (2018). Towards Neural Machine Translation for African Languages. *ArXiv:1811.05467 [Cs, Stat]*. https://arxiv.org/abs/1811.05467v1

ACAPS. (2019). *Lessons Learned: Needs Assessments in Cox's Bazar*. ACAPS. https://www.acaps.org/sites/acaps/files/products/files/20190404_acaps_npm_coxs_bazar_analysis_hub_needs_assessments_lessons_learned.pdf

AI Now Institute. (2018). *AI in 2018: A year in review*. Medium. https://medium.com/@AINowInstitute/ai-in-2018-a-year-in-review-8b161ead2b4e

Al-Amer, R., Ramjan, L., Glew, P., Darwish, M., & Salamonson, Y. (2016). Language translation challenges with Arabic speakers participating in qualitative research studies. *International Journal of Nursing Studies*, *54*, 150–157. https://doi.org/10.1016/j.ijnurstu.2015.04.010

Alzghool, M. (2009). *Investigating different models for cross-language information retrieval from automatic speech transcripts* [University of Ottawa]. https://ruor.uottawa.ca/handle/10393/29900

Amini, A., Soleimany, A., Schwarting, W., Bhatia, S., & Rus, D. (2019). Uncovering and Mitigating Algorithmic Bias through Learned Latent Structure. *AIES*. http://www.aies-conference.com/wp-content/papers/main/AIES-19_paper_220.pdf

Anokwa, Y., Hartung, C., & Brunette, W. (2009). Open source data collection in the developing world. *Computer*, *42*(10), 97–99. https://doi.org/10.1109/MC.2009.328

Ansari, A., & Petras, R. (2018). *Gamayun: The Language Equality Initiative*. Translators Without Borders. https://translatorswithoutborders.org/wp-content/uploads/2018/03/Gamayun-Language-Equality-Initiative-March-2018.pdf

Assal, H., Seng, J., Kurfess, F., Schwarz, E., & Pohl, K. (2011). Semantically-enhanced information extraction. *2011 Aerospace Conference*, 1–14. https://doi.org/10.1109/AERO.2011.5747547

Bansal, S., Kamper, H., Livescu, K., Lopez, A., & Goldwater, S. (2018). Low-resource speech-to-text translation. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, *2018-Septe*, 1298–1302. https://doi.org/10.21437/Interspeech.2018-1326

Bauer, J.-M., Akakpo, K., Enlund, M., & Passeri, S. (2013). Tracking Vulnerability in Real Time: Mobile Text for Food Security Surveys in Eastern Democratic Republic of Congo. *Africa Policy Journal*, *9*, 36. https://search.proquest.com/docview/1620000489

Bonino, F., Jean, I., & Knox Clarke, P. (2014). *Humanitarian Feedback Mechanisms: Research, Evidence and Guidance*. ALNAP/ODI. https://www.cdacollaborative.org/publication/humanitarian-feedback-mechanisms-research-evidence-and-guidanc/

Bowden, A., & Fox-Rushby, J. A. (2003). A systematic and critical review of the process of translation and adaptation of generic health-related quality of life measures in Africa, Asia, Eastern Europe, the Middle East, South America. *Social Science & Medicine*, *57*, 1289–1306. https://doi.org/10.1016/S0277-9536(02)00503-8

Brundage, M., Avin, S., Clark, J., Toner, H., Eckersley, P., Garfinkel, B., Dafoe, A., Scharre, P., Zeitzoff, T., Filar, B., Anderson, H., Roff, H., Allen, G. C., Steinhardt, J., Flynn, C., hÉigeartaigh, S. Ó., Beard, S., Belfield, H., Farquhar, S., … Amodei, D. (2018). *The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation*. http://arxiv.org/abs/1802.07228

Building Markets, & Orange Door Research. (2018). *What is the Point if Nothing Changes: Current Practices and Future Opportunities to Improve Remote Monitoring and Evaluation in Syria*. New York, NY: Building Markets.

Buolamwini, J. (2018). Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. *Proceedings of Machine Learning Research*, *81*, 1–15. http://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf

Clarke, P. K. (2018). *The State of the Humanitarian System 2018*. ALNAP/ODI. https://sohs.alnap.org/help-library/the-state-of-the-humanitarian-system-2018-full-report

CommonVoice. (2019). *Languages*. https://voice.mozilla.org/en/languages

Crawford, K., & Finn, M. (2015). The limits of crisis data: analytical and ethical challenges of using social and mobile data to understand disasters. *GeoJournal*, *80*(4), 491–502. https://doi.org/10.1007/s10708-014-9597-z

Cunningham, H., Tablan, V., Roberts, A., & Bontcheva, K. (2013). Getting More Out of Biomedical Documents with GATE's Full Lifecycle Open Source Text Analytics. *PLoS Computational Biology*, *9*(2), e1002854. https://doi.org/10.1371/journal.pcbi.1002854

Dette, R. (2018). Do No Digital Harm: Mitigating Technology Risks in Humanitarian Contexts. In S. Hostettler, S. Najih Besson, & J.-C. Bolay (Eds.), *Technologies for Development* (pp. 13–29). Springer. https://doi.org/10.1007/978-3-319-91068-0_2

Dette, R., Steets, J., & Sagmeister, E. (2016). *Technologies for monitoring in insecure environments*. Global Public Policy Institute. http://www.gppi.net/fileadmin/user_upload/media/pub/2016/SAVE__2016__Toolkit_on_Technologies_for_Monitoring_in_Insecure_Environments.pdf

Deutscher, G. (2010). *Through the Language Glass: Why the World Looks Different in Other*

Languages. Henry Hold and Company.

Emergency Data Science Workshop. (2018). *Natural Language Processing for Humanitarian Survey Work: Challenge Presented by International Rescue Committee*. https://emergencydatascience.org/challenge_nlp_surveys/

Facebook. (2017, September 14). *A New Center for Crisis Response on Facebook*. Menlo Park: Facebook. https://newsroom.fb.com/news/2017/09/a-new-center-for-crisis-response-on-facebook/

Furui, S., Nakamura, M., Ichiba, T., & Iwano, K. (n.d.). *Why is the Recognition of Spontaneous Speech so Hard?* Retrieved June 5, 2019, from http://www.furui.cs.titech.ac.jp

Future of Life Institute. (n.d.). *AI Policy Resources*. Retrieved March 5, 2019, from https://futureoflife.org/ai-policy-resources/

Ghazali, D. A., Guericolas, M., Thys, F., Sarasin, F., Arcos González, P., & Casalino, E. (2018). Climate change impacts on disaster and emergency medicine focusing on mitigation disruptive effects: An international perspective. *International Journal of Environmental Research and Public Health*, *15*(7), 1–13. https://doi.org/10.3390/ijerph15071379

Gibson, D. G., Pereira, A., Farrenkopf, B. A., Labrique, A. B., Pariyo, G. W., & Hyder, A. A. (2017). Mobile Phone Surveys for Collecting Population-Level Estimates in Low- and Middle-Income Countries: A Literature Review. *Journal of Medical Internet Research*, *19*(5), e139. https://doi.org/10.2196/jmir.7428

Google Cloud. (n.d.). *Language support*. Retrieved April 20, 2019, from https://cloud.google.com/speech-to-text/docs/languages

Greenough, P. G., Chan, J. L., Meier, P., Bateman, L., & Dutta, S. (2009). Applied technologies in humanitarian assistance: Report of the 2009 applied technology working group.

Prehospital and Disaster Medicine, *24*(SUPPL.2), 2–5. https://doi.org/10.1017/S1049023X00021609

Gu, J., Wang, Y., Chen, Y., Cho, K., & Li, V. O. K. (2018). *Meta-Learning for Low-Resource Neural Machine Translation*. http://arxiv.org/abs/1808.08437

Hannun, A., Case, C., Casper, J., Catanzaro, B., Diamos, G., Elsen, E., Prenger, R., Satheesh, S., Sengupta, S., Coates, A., & Ng, A. Y. (2014). *Deep Speech: Scaling up end-to-end speech recognition*. http://arxiv.org/abs/1412.5567

Harvard Humanitarian Initiative. (2011). *Disaster Relief 2.0: The Future of Information Sharing in Humanitarian Emergencies*. https://hhi.harvard.edu/sites/default/files/publications/disaster-relief-2.0.pdf

HHI and ICRC. (2018). *Engaging with people affected by armed conflicts and other situations of violence – Taking stock. Mapping trends. Looking ahead. Recommendations for humanitarian organizations and donors in the digital era.* https://www.icrc.org/en/publication/engaging-people-affected-armed-conflicts-and-other-siutations-violence

High-Level Expert Group on Artificial Intelligence. (2019a). *A definition of Artificial Intelligence: main capabilities and scientific disciplines*. https://ec.europa.eu/digital-single-market/en/news/definition-artificial-intelligence-main-capabilities-and-scientific-disciplines

High-Level Expert Group on Artificial Intelligence. (2019b). *Ethics guidelines for trustworthy AI*. https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=58477

Hilhorst, D., & Schmiemann, N. (2002). Development in Practice Humanitarian principles and organisational culture: Everyday practice in Médecins Sans Frontières-Holland. *Development in Practice*, *12*(3–4), 490–500. https://doi.org/10.1080/0961450220149834

Humanitarian Outcomes. (2021). *Aid Worker Security Database*. https://aidworkersecurity.org/incidents/report

Hunt, M., Pringle, J., Christen, M., Eckenwiler, L., Schwartz, L., & Davé, A. (2016). Ethics of emergent information and communication technology applications in humanitarian medical assistance. *International Health*, *8*(4), 239–245. https://doi.org/10.1093/inthealth/ihw028

IASC. (1999). *Protection of Internally Displaced Persons*. https://interagencystandingcommittee.org/system/files/legacy_files/protection_of_internally_displaced_persons_inter_agency_standing_committee_policy_paper_0.pdf

IASC. (2015). Multi-Sector Initial Rapid Assessment Guidance. In *Geneva: Inter-Agency Standing Committee*. https://interagencystandingcommittee.org/system/files/mira_2015_final.pdf

International Committee of the Red Cross. (2017). *Handbook on Data Protection in Humanitarian Action* (C. Kuner & M. Marelli (Eds.); 2017th ed.). International Committee of the Red Cross. https://shop.icrc.org/e-books/handbook-on-data-protection-in-humanitarian-action.html

ITU. (2018). *Measuring the Information Society Report 2018*. https://www.itu.int/en/ITU-D/Statistics/Documents/publications/misr2018/MISR-2018-Vol-1-E.pdf

Jacobsen, K. L. (2015). *The politics of humanitarian technology : good intentions, unintended consequences and insecurity*. Routledge.

Klein, G., Kim, Y., Deng, Y., Senellart, J., & Rush, A. M. (2017). *OpenNMT: Open-Source Toolkit for Neural Machine Translation*. http://arxiv.org/abs/1701.02810

Koehn, P., & Knowles, R. (2017). Six Challenges for Neural Machine Translation. *Proceedings Ofthe First Workshop on Neural Machine Translation*, 28–39.

https://doi.org/10.18653/v1/w17-3204

Kreutzer, T. (2009). *Internet and Online Media Usage on Mobile Phones among Low-Income Urban Youth in Cape Town*. Presented at International Communication Association Conference. http://tinokreutzer.org/mobile/InternetOnlineMediaUsage(ICA).pdf

Lamanna, C., Hachhethu, K., Chesterman, S., Singhal, G., Mwongela, B., Ng'endo, M., Passeri, S., Farhikhtah, A., Kadiyala, S., Bauer, J.-M., & Rosenstock, T. S. (2019). Strengths and limitations of computer assisted telephone interviews (CATI) for nutrition data collection in rural Kenya. *PLoS ONE*, *14*(1), e0210050. https://doi.org/10.1371/journal.pone.0210050

Langhaug, L. F., Cheung, Y. B., Pascoe, S. J. S., Chirawu, P., Woelk, G., Hayes, R. J., & Cowan, F. M. (2011). How you ask really matters: Randomised comparison of four sexual behaviour questionnaire delivery modes in Zimbabwean youth. *Sexually Transmitted Infections*, *87*(2), 165–173. https://doi.org/10.1136/sti.2009.037374

Leo, B., Morello, R., Mellon, J., Peixoto, T., & Davenport, S. (2015). Do Mobile Phone Surveys Work in Poor Countries? *CGD Working Paper*, *398*. https://doi.org/10.2139/ssrn.2597885

Loper, E., & Bird, S. (2002, May 17). NLTK: The Natural Language Toolkit. *Proceedings of the ACL Workshop on Effective Tools and Methodologies for Teaching Natural Language Processing and Computational Linguistics*. http://arxiv.org/abs/cs/0205028

Maya Hasan, M. (2018). *The language lesson: What we've learned about communicating with Rohingya refugees*. https://translatorswithoutborders.org/wp-content/uploads/2018/12/TWB_Bangladesh_Comprehension_Study_Nov2018.pdf

McCallum, A. K. (2002). *MALLET: A Machine Learning for Language Toolkit*. http://mallet.cs.umass.edu/

McNamara, A., Smith, J., & Murphy-Hill, E. (2018). *Does ACM's code of ethics change ethical*

decision making in software development? 729–733. https://doi.org/10.1145/3236024.3264833

Meier, P. (2011). New information technologies and their impact on the humanitarian sector. *International Review of the Red Cross*, *93*(884). https://doi.org/10.1017/S1816383112000318

Meier, P. (2014). *Digital humanitarians : how big data is changing the face of humanitarian response*. CRC Press.

Microsoft. (2018). *Microsoft reaches a historic milestone, using AI to match human performance in translating news from Chinese to English - Microsoft News Center Hong Kong*. https://news.microsoft.com/en-hk/2018/03/15/microsoft-reaches-a-historic-milestone-using-ai-to-match-human-performance-in-translating-news-from-chinese-to-english/

Mihalcea, R. (2004). Graph-based ranking algorithms for sentence extraction, applied to text summarization. *Proceedings of the ACL 2004 on Interactive Poster and Demonstration Sessions*. https://doi.org/10.3115/1219044.1219064

Mock, N., Morrow, N., & Papendieck, A. (2013). From complexity to food security decision-support: Novel methods of assessment and their role in enhancing the timeliness and relevance of food and nutrition security information. *Global Food Security*, *2*(1), 41–49. https://doi.org/10.1016/j.gfs.2012.11.007

Molnar, P., & Gill, L. (2018). *Bots at the Gate: A human rights analysis of automated decision-making in Canada's immigration and refugee system*. University of Toronto. https://ihrp.law.utoronto.ca/

Montreal declaration. (2018). *Montreal declaration for a responsible development of artificial intelligence 2018*. https://www.montrealdeclaration-responsibleai.com/the-declaration

Morrow, N., Mock, N., Bauer, J. M., & Browning, J. (2016). Knowing just in time: Use cases for mobile surveys in the humanitarian world. *Procedia Engineering*, *159*(June), 210–216. https://doi.org/10.1016/j.proeng.2016.08.163

OCHA. (2015). World Humanitarian Data and Trends 2015. In *United Nations Office for the Coordination of Humanitarian Affairs*. https://www.unocha.org/sites/dms/Documents/WHDT2015_2Dec.pdf

OCHA. (2018). *Global Humanitarian Overview 2019*. Office for the Coordination of Humanitarian Affairs. https://www.unocha.org/sites/unocha/files/GHO2019.pdf

Recommendation of the Council on Artificial Intelligence, Pub. L. No. OECD/LEGAL/0449 (2019). https://doi.org/10.1787/eedfee77-en

Paul, A., Jolley, C., & Anthony, A. (2018). *Reflecting the Past, Shaping the Future: Making AI Work for International Development*. USAID. https://www.usaid.gov/digital-development/machine-learning/AI-ML-in-development

Pham, P. N., & Vinck, P. (2012). Technology fusion and their implications for conflict early warning systems, public health, and human rights. *Health and Human Rights*, *14*(2), 106–117. http://www.jstor.org/stable/healhumarigh.14.2.106

Pictet, J. (1979). *The fundamental principles of the Red Cross*. http://www.ifrc.org/PageFiles/40669/Pictet Commentary.pdf

Poole, D., Latonero, M., & Berens, J. (2017). *Refugee Connectivity: A Survey of Mobile Phones, Mental Health, and Privacy at a Syrian Refugee Camp in Greece*. Cambridge, MA: Harvard Humanitarian Initiative and New York, NY: Data & Society Research Institute.

Raymond, N. A., & Card, B. (2015). *Applying Humanitarian Principles to Current Uses of Information Communication Technologies: Gaps in Doctrine and Challenges to Practice*.

Harvard Humanitarian Initiative. https://hhi.harvard.edu/sites/default/files/publications/signal_program_humanitarian_principles_white_paper.pdf

Robinson, A., & Obrecht, A. (2016). *Using mobile voice technology to improve the collection of food security data: WFP's mobile Vulnerability Analysis and Mapping*. ODI/ALNAP. https://www.alnap.org/system/files/content/resource/files/main/alnap-wfp-mvam-case-study-2016.pdf

Russell, S. J., & Norvig, P. (2009). *Artificial Intelligence: A Modern Approach* (3rd editio). Prentice Hall.

Sajad Mousavi, S., Schukat, M., & Howley, E. (2018). Deep Reinforcement Learning: An Overview. *SAI Intelligent Systems Conference (IntelliSys)*. https://arxiv.org/pdf/1806.08894.pdf

Sandvik, K. B., Gabrielsen Jumbert, M., Karlsrud, J., & Kaufmann, M. (2014). Humanitarian technology: a critical research agenda. *International Review of the Red Cross*, *96*(893), 219–242. https://doi.org/10.1017/s1816383114000344

Sandvik, K. B., Jacobsen, K. L., & McDonald, S. M. (2017). Do no harm: A taxonomy of the challenges of humanitarian experimentation. *International Review of the Red Cross*, *99*(904), 319–344. https://doi.org/10.1017/S181638311700042X

Sandvik, K. B., & Raymond, N. A. (2017). Beyond the Protective Effect: Towards a Theory of Harm for Information Communication Technologies in Mass Atrocity Response. *Genocide Studies and Prevention: An International Journal*, *11*(1), 9–24. https://doi.org/10.5038/1911-9933.11.1.1454

Scott-Smith, T. (2016). Humanitarian neophilia: the 'innovation turn' and its implications. *Third*

World Quarterly, *37*(12), 2229–2251. https://doi.org/10.1080/01436597.2016.1176856

Singh, S. (2018). *Natural Language Processing for Information Extraction*. http://arxiv.org/abs/1807.02383

Steets, J., Sagmeister, E., & Ruppert, L. (2016). *Eyes and Ears on the Ground: Monitoring aid in insecure environments*. Global Public Policy Institute. http://www.gppi.net/fileadmin/user_upload/media/pub/2016/SAVE__2016__Monitoring_aid_in_insecure_environments.pdf

Tatman, R. (2017). Gender and Dialect Bias in YouTube's Automatic Captions. *Proc. of the First Workshop on Ethics in Natural Language Processing*, 53–59. http://www.aclweb.org/anthology/W17-1606

The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems. (2019). *Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems* (First edit). https://doi.org/10.1109/MCS.2018.2810458

Translators Without Borders. (2018a). *The words between us: How well do enumerators understand the terminology used in humanitarian surveys? A study from northeast Nigeria*. https://translatorswithoutborders.org/wp-content/uploads/2018/12/TWB_Nigeria_EnumeratorComprehension_Nov2018-1.pdf

Translators Without Borders. (2018b, April 19). *Communications Dashboard: Internally Displaced People in Northeast Nigeria*. https://translatorswithoutborders.org/communications-dashboard-internally-displaced-people-in-north-east-nigeria/

Translators Without Borders. (2019). *Language mapping: Putting communication needs on the map*. https://translatorswithoutborders.org/language-mapping/

UNHCR. (2016). *Connecting Refugees: How Internet and Mobile Connectivity can Improve Refugee Well-Being and Transform Humanitarian Action*. http://www.unhcr.org/5770d43c4.pdf

UNHCR. (2018). *Global Trends 2017*. https://www.unhcr.org/5b27be547.pdf

Van der Windt, P., & Humphreys, M. (2016). Crowdseeding in Eastern Congo. *Journal of Conflict Resolution*, *60*(4), 748–781. https://doi.org/10.1177/0022002714553104

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). *Attention Is All You Need*. https://arxiv.org/pdf/1706.03762.pdf

Vincent, J. (2019). *AI researchers debate the ethics of sharing potentially harmful programs*. The Verge. https://www.theverge.com/2019/2/21/18234500/ai-ethics-debate-researchers-harmful-programs-openai

Weiss, R. J., Chorowski, J., Jaitly, N., Wu, Y., & Chen, Z. (2017). Sequence-to-sequence models can directly translate foreign speech. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, *2017-Augus*, 2625–2629. https://doi.org/10.21437/Interspeech.2017-503

WHO. (2007). *Risk reduction and emergency preparedness*. World Health Organization. https://apps.who.int/iris/bitstream/handle/10665/43736/9789241595896_eng.pdf

Wood, G. (2018, April). The Refugee Detectives. *The Atlantic*. https://www.theatlantic.com/magazine/archive/2018/04/the-refugee-detectives/554090/

3.7. Chapter 3 - Appendix 1

Statement of Authorship for Inclusion in Dissertation

Title of Study: Improving Humanitarian Needs Assessments through Natural Language Processing

Publication Status: Published

Publication Details: Kreutzer, T., Vinck, P., Pham, P., An, A., Appel, L., Deluca, E., Tang, G., Alzghool, M., Hachhethu, K., Morris, B., Crowley, J., & Orbinski, J. (2020). Improving Humanitarian Needs Assessments through Natural Language Processing. *IBM Journal of Research and Development*, *64*(1). https://doi.org/10.1147/JRD.2019.2947014

Student Name: Tino Kreutzer

Supervisor Name: James Orbinski

Contribution to Paper: Performed the study, wrote the initial manuscript, integrated contributions and feedback, and finalized the manuscript.

Signature: **Date**: 15 February 2023

Supervisor Confirmation

By signing the Statement of Authorship, I am certifying that the student made a substantial contribution to the publication, and that the description described above is accurate.

Signature: **Date**: 15 February 2023

Chapter 4 Overcoming Blind Spots: Constraints and Solutions Related to Qualitative Interview Methods in Humanitarian Needs Assessments

Abstract

Humanitarian needs assessments (HNAs) are used to prioritize and inform the provision of humanitarian assistance by humanitarian actors. Even so, empirical data about contemporary methodological HNA practices are lacking. This article aims to explore current approaches to conducting HNA, constraints related to different interview methods, with particular attention to qualitative methods and solutions for improving the use of qualitative interview methods in HNAs. In-depth interviews were conducted with international key informants (KIs) with prior experience managing HNA data collection. 23 KIs from 13 countries participated, representing 17 humanitarian organizations. Results from the transcribed semi-structured interviews resulted in 47 themes, which were organized into five Thematic Groups: 1) Prevailing approaches to data collection in HNA; 2) constraints related to quantitative interview methods; 3) constraints related to qualitative interview methods; 4) challenges specific to pre-coded open-ended questions; and 5) respondents' recommendations for improving the role of qualitative interview methods in HNA. Respondents pointed to a complex operational environment that favors quantitative methods in HNA, even though qualitative data is widely seen as crucial for understanding population needs. Principle findings show that KIs suggested acknowledging the limitations of quantitative data, investing in more qualitative skills, piloting mixed-methods hybrid surveys, as well as employing natural language processing, a type of artificial intelligence, for automatic speech recognition, machine translation, and content analysis.

Overcoming Blind Spots: Constraints and Solutions Related to Qualitative Interview Methods in Humanitarian Needs Assessments

In 2021, conflicts and disasters have led to more than 274 million people requiring humanitarian assistance (OCHA, 2021a). Organizations involved in providing humanitarian assistance and responding to global health emergencies work under strenuous circumstances and with limited funding to provide life-saving aid to those with the greatest need. In 2021, the cost of humanitarian assistance amounted to US $ 41 billion (OCHA, 2021a).

Humanitarian needs assessments (HNAs) provide evidence of critical population needs in order to decide on the most appropriate interventions, to prioritize limited funding, and to create a baseline for measuring the impact of humanitarian programs through concrete evidence (Banatvala, 2000; Cosgrave, 2009; Gerdin et al., 2014; Redmond, 2005; Spiegel et al., 2001). Affected individuals and communities are an indispensable source of primary first-hand information about the needs, preferences, and local resources available for effective humanitarian assistance. Primary data obtained from affected people through interviews and other methods can more precisely match available humanitarian capacity on the one hand and urgent population needs on the other.

Humanitarian assistance is understood here to refer to coordinated actions that save lives, alleviate suffering, and maintain human dignity during and after human-made crises, and disasters caused by natural hazards. It also includes “protection,” which “encompasses all activities aimed at obtaining full respect for the rights of the individual in accordance with the letter and the spirit of the relevant bodies of law” (IASC, 1999, p. 4). Humanitarian needs assessments are understood here as the “set of activities necessary to understand a given situation, [which] entails the collection, updating and analysis of data pertaining to the

population of concern (needs, capacities, resources, etc.), as well as the state of infrastructure and general socio-economic conditions in a given location/area" (UNHCR, 2006, p. 4). Such HNA may include many forms of primary data collection, either by a single organization or in combination with others (often referred to as coordinated, joint, or harmonized assessments depending on the level of collaboration). In this study, structured interviews refer to data collection using standardized questionnaires in which primarily closed-ended questions are asked in a pre-determined order. Semi-structured interviews refer to data collection based on a list of primarily open-ended questions (OEQ) that are used to stimulate and guide discussions (ACAPS, 2014; Creswell, 2014).

Three types of HNA can be distinguished based on the type of crisis, phase of the response, objectives of the humanitarian organizations involved, decisions by organizations or donors to inform, and time available for collecting information—as well as the kinds of data collection methods that are appropriate. This is summarized in Table 8 (UNHCR, 2017). Following a sudden-onset disaster (such as an earthquake or a rapid expansion of an armed conflict), initial assessments are conducted during the first two weeks, often referred to as Phase 1. These typically involve secondary data review (SDR), direct observations of damage, as well as data that can be collected remotely (e.g., flyovers, satellite imagery) and help to define the scale and severity of the crisis. Rapid assessments conducted during the third and fourth weeks, often referred to as Phase 2, typically include the collection of primary data from key informant interviews (KII) or community group discussions, depending on factors such as access to the population, resources, staffing, and safety considerations (IASC, 2015). Rapid assessments are intended to inform the initial planning of humanitarian assistance by establishing the severity of needs and the key priorities of the affected population. In-depth assessments (conducted after the

first four weeks following a rapid onset disaster, also known as Phase 3) typically include methods that require more time and resources and therefore often include detailed representative surveys, focus group discussions, key informant interviews (KII), and other methods. In-depth assessments may include many rounds of data collection depending on the scale, complexity, and duration of the crisis, and on the capacity to serially collect such data. In protracted humanitarian crises, understood here as armed conflicts that persist over long periods of time (ICRC, 2016), multi-sector needs assessments (MSNAs) are often conducted annually or at similar intervals to inform long-term programming or to prioritize donor funding. Slow-onset or cyclical crises may not require initial or rapid assessments, whereas some emergencies can include several waves of HNA from all three phases.

Table 8

Phases and Types of Humanitarian Needs Assessments

	Phase 1	**Phase 2**	**Phase 3**
Assessment Type	Initial assessment	Rapid assessment	In-depth assessment
Timeframe	First 2 weeks	2-4 weeks	1-4 months; cyclical
Example objectives	Define scale and severity of the crisis Define access constraints	Assess severity of needs Establish key priorities Identify information gaps	Define and quantify needs by sector Capture representative views of affected populations Establish baseline for needs and response monitoring
Example of decisions to inform	Emergency funding appeals	Initial planning of humanitarian response	Inform detailed planning and scope of humanitarian programming
Typical designs and data collection techniques	SDR Remote primary data collection (e.g., flyover, satellite) Direct observations	SDR Community group discussions KII at the community level	SDR Surveys (households/individuals) Focus group discussions KII
Outputs	Secondary data review report; emergency call or appeal; situation report; rapid decisions on urgent aid/staff deployments	Overview report by sector, geographic locations, and population sub-groups; prioritization of needs; short-term program decisions	Detailed assessment reports; long-term programming decision for organizations/clusters Humanitarian Needs Overview and Humanitarian Response Plan

HNAs are conducted through a combination of quantitative and qualitative methods. Purely quantitative methods may include collecting naturally occurring numerical data (e.g., the number of children in a household, income levels, or mid-upper arm circumference

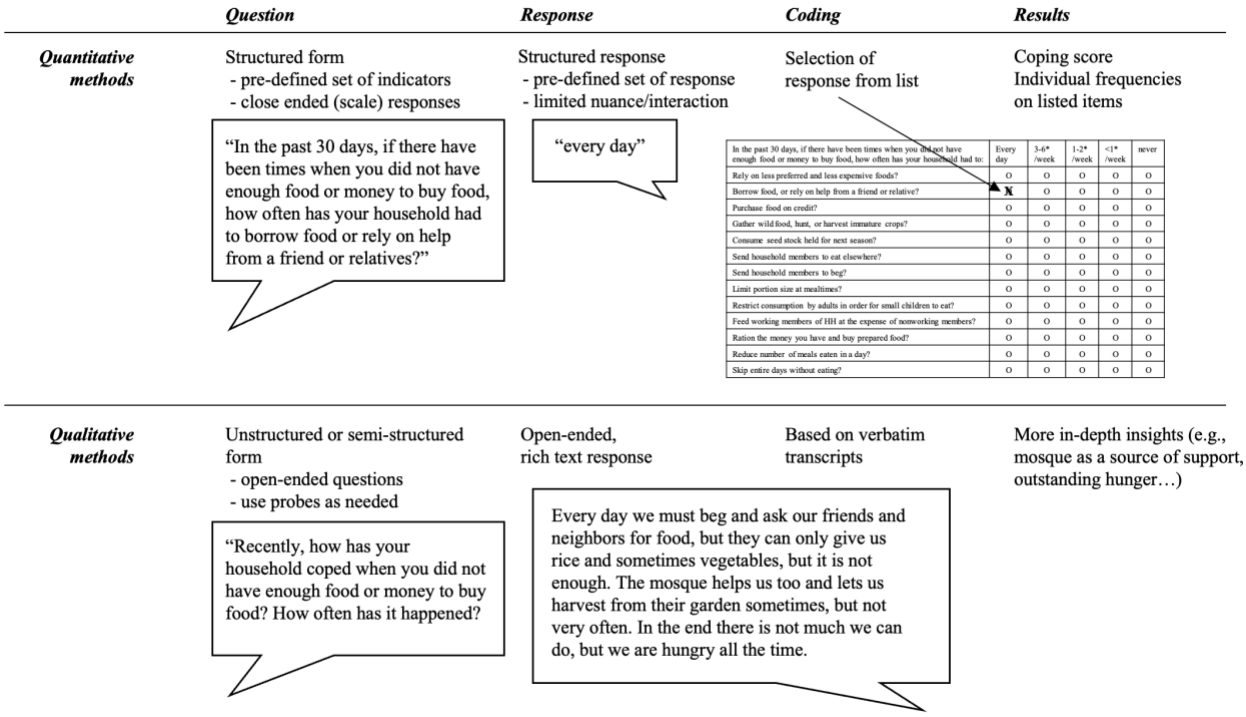
measurements), for example through surveys of randomly selected households. HNA data from structured interviews are commonly collected by entering responses into handheld computer-assisted personal interviewing (CAPI) tools. KoboToolbox (Kobo, 2022), a CAPI tool created with support of the UN Office for the Coordination of Humanitarian Affairs (OCHA) and in partnership with the UN High Commissioner for Refugees (UNHCR), is used by the majority of humanitarian organizations for HNA and other purposes (Building Markets & Orange Door Research, 2018). Qualitative methods are used to capture more nuanced insights, often by recording responses to OEQ in audio or text format for later analysis, for example through key informant interviews or focus group discussions. Qualitative methods are usually preferred during rapid assessments, whereas a mixed-methods approach using both qualitative and quantitative data collection methods is often considered optimal for informing relevant humanitarian assistance during in-depth assessments (OCHA, 2021b; UNHCR, 2017). For example, semi-structured key informant interviews may establish the presence of local challenges to accessing assistance, a household survey may then assess the presence of these issues across geographic areas through a representative sample, and focus group discussions may then be used to understand and contextualize the trends identified or discovered during the survey.

However, as described in our earlier study (Kreutzer et al., 2020), numerous challenges prevent organizations from using qualitative methods at a larger scale as part of HNA. In practice, current HNA approaches often require interviewers to convert complex responses to OEQ in surveys or KII into simplified quantitative data or short notes. This is typically done with the use of structured questionnaires by having interviewers code responses in the field, or by writing down short sentences that are then coded during analysis (see Figure 6). Furthermore, the

diversity of languages and dialects common in many humanitarian crises can make collecting and analyzing qualitative data both difficult and costly (ACAPS, 2019; Al-Amer et al., 2016; Bowden & Fox-Rushby, 2003; Deutscher, 2010; Translators Without Borders, 2018a, 2018b, 2019).

Figure 6

Schematic Comparison Between Quantitative and Qualitative Interview Methods



Efforts to improve the quality and usability of HNA have included several initiatives in recent years. The 2016 Humanitarian Grand Bargain included several workstreams aimed at improving HNA (WHS, 2016), which in turn led to a review of 39 different HNA analytical frameworks (Chataigner, 2017), a method of scoring and comparing the quality of HNAs (GPPI & INSPIRE Consortium, 2019; Okular Analytics, 2021), as well as the creation of the Joint Intersectoral Analysis Framework (JIAF) in 2021 (OCHA, 2021b). Most of this work has focused primarily on improving the use of quantitative household-level data. To date, however, there has been very little focus on the challenges of using qualitative methods in HNA.

This research aimed, therefore, to explore 1) current approaches to conducting HNA, 2) constraints related to different interview methods, particularly qualitative ones, and 3) solutions for improving the use of qualitative interview methods in HNAs. This study presents data from interviews with 23 key informants (KIs), each with intimate knowledge of designing, conducting, and analyzing HNAs across a wide range of humanitarian organizations.

4.1. Methods

A qualitative study with semi-structured individual interviews with KIs was conducted between 2020 and 2021. Ethics approval for the study was granted by the Office of Research Ethics at York University, Canada (certificate # STU 2020-092). Informed consent was obtained from all participants.

4.1.1. Sampling and Recruitment

KIs were selected using exponential discriminative snowball sampling, starting with existing KIs known to the author. KIs had to be currently working full time in humanitarian assistance and possess prior experience managing HNA data collection, as well as be able to fluently speak English, French, German, or Spanish as these were languages spoken by the researcher. After the first 10 KIs were interviewed, additional respondents were sought to maximize heterogeneity of sampled organizations, such as United Nations (UN) agencies, the Red Cross/Red Crescent Movement, as well as international and national or local nongovernmental organizations (NGOs). Several attempts were made to include representatives of national or local NGOs from countries affected by humanitarian crises, but these were unsuccessful. Although KIs were (in some cases) intentionally sampled as representing their parent organizations to enable results to be reflective of the broad humanitarian sector, their

views and experiences do not represent official statements from the respective organizations. In fact, because many KIs have worked for several humanitarian organizations, their views were often more reflective of their overall experience rather than related to a specific humanitarian employer.

KIs were invited to participate by means of a personal email that included an outline of the study purpose, a link to the electronic consent form in English, and a way to self-schedule the interview based on the most convenient date and time. The consent form allowed KIs to consent to (or refuse) having an audio recording of the interview created and to having their responses quoted anonymously.

4.1.2. Interview Instrument Development

A semi-structured interview guide (available in Chapter 4 - Appendix 3) was iteratively developed to explore current practices and perceptions of:

1. Prevailing approaches to conducting HNAs to inform operational decision-making, challenges in understanding population needs, and the impact of the COVID-19 pandemic on conducting HNAs
2. The role of quantitative and qualitative interview methods used in HNAs during different stages of humanitarian assistance, including methods of processing OEQs and the role of audio recordings, as well as suggestions for addressing existing challenges
3. Ethical concerns and data protection approaches related to processing HNAs
4. The present study focuses on the issues covered in points one and two; a separate study will discuss the findings from point three.

4.1.3. Data Collection

Pilot Phase. The initial interview guide was tested with three KIs from three countries, representing a UN Agency and two international NGOs. Revisions were made to the interview questions following interim data analysis, and the final data collection instrument was developed.

Interviews. Data collection took place between November 2020 and May 2021. All interviews were conducted in English by the author (a social scientist based in Canada) via the Zoom video conferencing platform (Gray et al., 2020). The audio was recorded for verbatim transcription, supported by notes. Recruitment and data collection were completed (n = 23) at the point when no new concepts were found to emerge.

Data Processing and Analysis. Interview transcripts were analyzed in MAXQDA 2020 (*MAXQDA 2020*, 2019) using a grounded theory approach through which theoretical insights are identified from the data in an iterative process of reading and coding (Belgrave & Seide, 2018; Chun Tie et al., 2019; Corbin & Strauss, 2015). Initial codes were generated by applying an open coding approach line by line to the first three transcripts following the pilot phase. The initial codes were reviewed and organized before being applied iteratively to the other 20 interview transcripts. Additional codes were developed while other categories were collapsed through constant comparative analysis of previously coded transcript segments. Throughout this process, memos were created to assist initial analysis and contextualization of the results. In the second analysis stage, codes were analyzed thematically for the topics covered in this manuscript. Quotations included in this study are illustrative of common perceptions and experiences reported by the KIs.

To report on the respondent characteristics, countries were categorized based on the self-reported country of residence, rather than country of birth, which was done in order to reflect the

current work and cultural environment shaping the opinions of KIs. Countries were then grouped by geographic region, using the World Bank's classification scheme (World Bank, 2020). Great care was taken to protect the anonymity of respondents. Quotations from KIs are therefore not attributed to their organization or country of residency. Likewise, this study does not report respondents' country of birth separately as this information, together with the list of organizations and the country of residency, could lead to the identification of some individuals. This manuscript follows the Standards for Reporting Qualitative Research checklist (O'Brien et al., 2014) (SRQR, see Chapter 4 - Appendix 1).

4.2. Results

4.2.1. Sample Characteristics

23 KIs (female: n = 14) from 13 countries participated, collectively representing 17 humanitarian organizations (see Table 9). Most participants were residents of countries in Europe (n = 11) and North America (n = 5). Just over half (n = 12) of respondents represented a non-governmental organization, three were from an organization of the Red Cross/Red Crescent movement, six represented a UN organization, one KI was from the Global Health Cluster, and one was an independent consultant. Organizations listed in Table 9 represent the current affiliation at the time of interview; however, six KIs changed organizations within three months preceding the interview, while in most cases, respondents also drew on experiences and perceptions based on their work at multiple organizations prior to their current humanitarian employer. These recent employment changes also led to the inclusion of two additional KIs who had recently moved to the International Office for Migration (IOM). The mean duration of interviews was 89 minutes (SD = 26; range = 46–144 min).

Table 9

Sample Distribution across Countries of Residence and Organizations Represented (N = 23)

Countries	Count	Organizations	Count
East Asia & Pacific	2	**NGO**	12
Australia	2	ACAPS*	1
Europe & Central Asia	11	ACF (Action Contre la Faim)	1
France	1	DRC (Danish Refugee Council)	1
Italy	1	iMMAP*	1
Switzerland	5	IRC (International Refugee Committee)	2
United Kingdom	4	JIPS (Joint IDP Profiling Service)	1
Latin America & Caribbean	1	Mercy Corps	1
Colombia	1	MSF (Médecins Sans Frontières)	1
Middle East & North Africa	2	REACH/IMPACT Initiatives*	1
Jordan	1	Save the Children	2
Syrian Arab Republic	1	**Red Cross/Red Crescent Movement**	3
North America	5	ICRC (International Committee of the Red Cross)	1
Canada	1	IFRC (International Federation of the Red Cross and Red Crescent Societies)	2
United States	4	**United Nations**	6
South Asia	1	IOM	3
India	1	OCHA	1
Sub-Saharan Africa	1	UNHCR	1
Ethiopia	1	WFP (World Food Programme)	1
		Other	2
		WHO Global Health Cluster	1
		Independent consultant	1

* ACAPS, iMMAP, and REACH/IMPACT Initiatives are the official organizations' names and are not or no longer considered acronyms

4.2.2. Results Framework

Qualitative analysis results from the transcribed interviews resulted in 47 themes, which were organized into five Thematic Groups: 1) prevailing approaches to data collection in HNA (12 themes); 2) constraints related to quantitative interview methods (9 themes); 3) constraints

related to qualitative interview methods (14 themes); 4) challenges specific to pre-coded OEQs (4 themes); and 5) suggestions for improving the role of qualitative interview methods in HNA (8 themes). The following sections summarize the results of the most salient themes. The summary table of all themes identified is described in Table 10; an extended list of themes, including descriptions and illustrative quotations for each of the identified and uncategorized 47 themes, is available in Chapter 4 - Appendix 2.

Table 10

Overview of Themes by Thematic Group and Theme

#	Theme
Thematic Group 1: Prevailing approaches to data collection in HNA	
1.1	Household surveys are the default assessment method for most organizations
1.2	Key informant interviews are predominantly quantitative
1.3	The role of quantitative data has increased due to KoboToolbox's widespread usage
1.4	OEQ are considered crucial for understanding complex subjects, but mostly used with pre-coded options
1.5	There is confusion about what constitutes qualitative and quantitative information
1.6	Uncoded OEQ can be crucial for filling in information gaps
1.7	Paper and Excel are often preferred over CAPI tools for taking long notes in KII
1.8	Donors are exerting pressure in favor of quantitative methods, particularly MSNA
1.9	Some organizations are biased in favor of quantitative methods
1.10	Qualitative methods are critical for answering "why" and "how" questions
1.11	Qualitative and quantitative methods are seen as complimentary
1.12	Qualitative skills are slowly increasing
Thematic Group 2: Constraints related to quantitative interview methods	
2.1	Some information by respondents is never recorded
2.2	For some sectors, household surveys are not the best method to obtain important information
2.3	Some groups are suspicious of quantitative assessments, particularly remote survey methods
2.4	Structured KII are overused
2.5	A lot of quantitative data is contradictory or is not explained
2.6	Limited access to affected population can bias assessment results
2.7	Household surveys can put respondents and staff in danger
2.8	Representative household surveys are not appropriate during Phases 1 and 2
2.9	Questionnaires are not contextualized enough
Thematic Group 3: Constraints related to qualitative interview methods	
Cluster 3.1: Capacity limitations	

#	Theme
3.1.1	Qualitative analysis is too time consuming for emergency contexts
3.1.2	There are not enough staff with expertise in qualitative methods
3.1.3	Interviewer training is much too short
Cluster 3.2: Notetaking / transcript challenges	
3.2.1	Manually created notes are often of poor quality
3.2.2	Verbatim transcripts of responses are only rarely created because audio is almost never recorded
3.2.3	Audio files can create data protection risks or make respondents more reticent
Cluster 3.3: Language issues	
3.3.1	Significant language gaps hamper qualitative data collection and analysis
3.3.2	Notes are rarely translated properly
3.3.3	Meaning is lost because analysis is not done by or with local staff
Cluster 3.4: Systemic challenges	
3.4.1	Qualitative and quantitative methods are often not integrated
3.4.2	There are systematic biases against qualitative data
3.4.3	The trend to collect data remotely reduces the ability to collect qualitative data
3.4.4	Donors' short funding cycles don't allow for qualitative methods
3.4.5	The push for more secondary data analysis and automation cannot replace qualitative primary data
Thematic Group 4: Challenges specific to pre-coded open-ended questions	
4.1	Pre-coded responses to open-ended questions in questionnaires can lead to false data
4.2	OEQ are considered an obstacle to obtaining large samples
4.3	Most text responses to OEQ are not properly analyzed
4.4	Interviewers may not choose the correct responses among pre-coded options
Thematic Group 5: Suggestions for improving the role of qualitative interview methods in HNA	
5.1	More funding is needed for longer training sessions
5.2	Ensure assessments teams have qualitative expertise and speak local languages
5.3	MSNA should contain more qualitative methods
5.4	Conduct hybrid surveys with a mix of quantitative and qualitative methods
5.5	More audio recordings would be useful to capture all details
5.6	Automatic transcription would allow more widespread qualitative data collection
5.7	Machine translation would be helpful for systematically translating transcripts or notes
5.8	Automatic coding would be helpful to speed up analysis

4.2.3. Thematic Group 1: Prevailing Approaches to Data Collection in HNA

For most KI organizations, household surveys and structured KIIs are the most common methods of collecting data in HNAs. These surveys were characterized as "quick shots" that

allow a rapid cycle between identifying an information need, collecting data, analyzing the data, and interpreting the results. The adoption of KoboToolbox was seen as an important factor behind the relative availability and importance of quantitative data in HNAs. All but one respondent mentioned that the tool was used in their organization, saying "I think we are the only humanitarian organization that doesn't use Kobo" (KI #18). Some respondents pointed out that prior to the widespread adoption of CAPI tools, assessments tended to be more qualitative in nature.

> But, you know, Kobo's been set up in such a way where it's so easy to get large volumes of data. . . . That day, you can do a survey and, in the evening, get back—or even in the car on the way back from the site, you can start to look at the data almost immediately. It's amazing. (KI #08)

In many cases, respondents reported that, ideally, household surveys should be used to complement qualitative methods. This complementarity was seen by most respondents as crucial for finding answers to deeper questions, with one noting that qualitative methods "really enhance the understanding that you might get from a purely quantitative viewpoint" (KI #12). Some respondents saw the role of qualitative methods as restricted to "coloring in the picture" (KI #03) by answering "how" and "why" questions, while quantitative methods should be used for answering high level "what" questions about population needs. However, others pointed out that some issues cannot be captured through quantitative means and would be overlooked without qualitative methods.

Most KIIs are conducted with fully structured questionnaires (sometimes including OEQs within a structured instrument). During Phases 1 and 2, KIIs with structured interview guides were mentioned to be often the first—or only—primary data collected for establishing community-level baselines for population size and key issues. Many respondents reported using

KoboToolbox or pre-formatted Excel tables to facilitate rapid data collection and analysis. Semi-structured KIIs are also conducted, but much less frequently. These practices fall short of the traditional understanding of this qualitative research method, "what an academic institution would consider key informant interviews" (KI #05).

Both quantitative and qualitative approaches use OEQs in interviews. KIs differentiated between two kinds of OEQ: 1) *Pre-coded OEQs* in structured surveys are programmed into CAPI instruments as multiple-choice questions (often with an "other" response to allow text entry for an unlisted choice); 2) *Uncoded OEQs*, used mostly in semi-structured KIIs, require the coding step to take place during the analysis stage. Text notes based on uncoded OEQs are often written on paper or typed into structured Excel tables (with one column per question), whereas CAPI tools are less commonly used. The distinction between qualitative and quantitative methods with regard to OEQs can lead to confusion. Some KIs considered pre-coded OEQs to be a shortcut method for capturing qualitative information but in a quantitative way. One KI argued that asking a family about their highest priorities during a household survey is not quantitative per se:

> That's qualitative for you and your household. But because we collect it from several thousand people and we present it in a graph, we make it into a quantitative thing. But it's still a qualitative piece of information. (KI #02)

Some respondents considered uncoded OEQs in semi-structured KIIs to be ideal for making best use of the respondent's time and filling in crucial information gaps. While they are not used widely, several respondents noted that uncoded OEQs have the potential to measure more useful details that are lost with pre-coded choices—particularly to measure concepts such as the "impact" of humanitarian assistance:

> People might speak more to their unique situation. And I think that if we had some way to sort of analyze it at that level, we would just be able to speak more to our impact. (KI #03)

Several KIs reported that in their experience there has been a strong push from some donor governments to increase the amount of quantitative data from household surveys across all humanitarian crises. They cited MSNAs in particular, which would enable donors to compare the situation in different crises and aid them in prioritization of funding. Several respondents involved in policy discussions linked this trend to the 2016 Grand Bargain commitments and the creation of the JIAF in 2021, which call for a more standardized approach of conducting HNAs.

Another driver for the increase in quantitative methods was seen to originate within humanitarian organizations, such as household surveys as a requirement of accessing internal funding, or for fine-tuning to population needs after an intervention had already been decided on. The most cited reason, however, was that people in leadership positions are often biased towards quantitative data:

> Management sees qualitative data as something that's easy to be cluttered if you want to have an easy approach. But I think the teams always rely only on household assessment data, because that is more scientific. That's, of course, the term that's often being used. (KI #18)

IOM and the World Food Programme (WFP) were identified by some respondents as fueling this increase in quantitative methods; both organizations are known for rapidly collecting and visualizing quantitative data, through the Displacement Tracking Matrix (DTM) and Vulnerability, Analysis and Mapping (VAM) programs (International Office for Migration, 2021; Morrow et al., 2016), respectively. KIs noted that the success of DTM and VAM had

increased expectations for other organizations to adopt similar approaches for rapidly collecting and visualizing more quantitative data.

Despite this growth in quantitative methods, some respondents described a growing demand for qualitative information in the humanitarian sector. Much of this growth was taking place in teams working in the related fields of accountability to affected people and community engagement (Global Health Cluster, 2017; International Committee of the Red Cross, 2019). However, the qualitative work in these areas was said to be done most often in parallel to HNA rather than in collaboration.

4.2.4. Thematic Group 2: Constraints Related to Quantitative Interview Methods

Many respondents argued that many organizations overuse structured questionnaires for KII, even if more qualitative approaches are possible. Similarly, many KIs also highlighted that, because of time constraints, standardized questionnaires are often used without being contextualized to the country and crisis in question, leading to potential omissions and biases. Several respondents argued that ideally, more contextualization and proper qualitative KIIs should be conducted during protracted crises. However, this is done only rarely—even once field site accessibility has improved—either out of habit or for reasons of convenience. As as result, HNAs based on quantitative data in protracted crises are often less reliable:

> [When conducting HNA] in the immediate aftermath of an event, we know our data is questionable. We are sacrificing immense quality, just to get it fast. . . . A few weeks down the road like this should not be the basis for everything. It's just the initial planning. So, we know it's questionable. Then the problem is when you're still using those methods when you no longer need to. And that happens far too often just because they're easy. (KI #05)

Data quality issues can also result from an abundance of quantitative data from multiple assessment rounds that can present contradictory results. Despite this wealth of data, respondents indicated that, sometimes, there is little actual understanding of the population. This appears to be the case particularly in refugee or IDP camp environments during protracted crises where successive in-depth quantitative assessments are easier to conduct:

> I'm drowning in all this secondary data [from quantitative surveys]. But I don't feel like I understand the population at all. And all I feel like I'm seeing is conflicting quantitative data that says this, many people need this, but then they're doing this behavior. (KI #10)

In some cases, such confusion may stem from overuse of the household survey as a device to measure all phenomena. Some KIs explained that for some humanitarian sectors (such as health or protection), household surveys are not an appropriate method for obtaining operationally useful information. This can even mean that many of the collected indicators are recognized as too unreliable, and therefore ignored—as in the case of the health sector where all questions but one were considered unreliable:

> [I am] trying to stop people from asking diagnostic questions in household surveys because people are not great at self-diagnosing. . . . We can't use it for anything, it's not operational. . . . The bulk of [what we could use] has been one question on "How far away is the nearest health facility in minutes?" (KI #05)

Several themes emerged related to difficulties in reaching populations during emergencies, especially due to logistics or limitations imposed by armed forces. KIs argued that such limitations lead to biased samples, while interviewers and respondents may be put in physical danger if seen as conducting household surveys. Some respondents argued that

conducting representative surveys during Phases 1 and 2 following an emergency was practically impossible and should therefore no longer be attempted: “Some people view [rapid needs assessments] as a household-level representative assessment that needs to be done in two weeks. Which is wrong, it can’t happen” (KI #14).

Some KIs noted that much of the potentially important contextual information given by respondents is systematically lost when quantitative survey instruments have nowhere to enter this data. One gave an example of a response to a standard yes/no question that would be useful to operational planning, if somehow recorded:

> If someone asks a question . . . “Do you have access to the markets?” “No . . . I got intercepted by military, I tried to do this, and the roads were like . . . . ” That’s interesting. And it’s not there. (KI #07)

Finally, certain populations are weary of quantitative interview methods due to their closed-ended nature that does not allow respondents to freely express themselves. This phenomenon, often referred to as survey fatigue, may be leading to potentially higher rates of refusal. Some KIs said this applied to specific population groups such as displaced Rohingya or Pacific Islanders, as well as possibly rural populations in general:

> They’ve reported it 100 times, they don’t like quantitative data collection, especially when they’re asking about people’s behavior and how they’re feeling and things like that. They don’t really quite understand why someone would want that information. They don’t find that consultative, they don’t find that any type of conversation. So, it’s hard to get any type of meaningful data if that’s not what they would like to communicate through. (KI #10)

4.2.5. Thematic Group 3: Constraints Related to Qualitative Interview Methods

Fourteen themes were identified related to factors that limit the utility of qualitative interview methods, which were further grouped into four clusters: capacity limitations (3 themes), notetaking/transcript challenges (3), language issues (3), and systemic challenges (5).

Cluster 3.1: Capacity Limitations. All KIs pointed out that organizations lack sufficiently skilled staff to collect and analyze qualitative data. Most organizations do not have enough staff with qualitative research skills, a theme that was discussed far more often than any other. However, there are significant differences between different institutions, with some being known for having far more qualitative expertise than others: "The organization in question often will tell you a lot: If it's [organization 1] I'll assume its proper qualitative. If it's [organization 2], I'm going to be really wary" (KI #05, names of humanitarian organizations redacted).

Respondents mentioned that in many organizations, assessments are increasingly being managed by information management specialists, who often have training in quantitative skills, such as database management or geographic information systems. One respondent in a senior information management role at a UN agency was unfamiliar with the term "qualitative methods." Instead, the KI admitted that they only rarely collected open-ended responses, "because it's a nightmare to analyze. . . . How do you recode it into something?" (KI #13).

Some larger NGO and UN organizations have dedicated emergency response teams with more qualitative experience who can deploy for short periods (often during Phase 2 or 3). But when these staff return home, long-term project staff are often unable to continue or repeat qualitative assessments:

> I think it's often easier with an emergency team, because you have . . . a smaller group [of] people, you train them. And usually, they're the people that are coming with quite a lot of experience. . . . Whereas at

> country program level . . . it's not necessarily those people with the same level of experience. (KI #12)

Proper analysis of qualitative data was considered too time consuming for emergency contexts where results need to be established quickly. One respondent argued that during sudden-onset emergencies, field teams should not even attempt using qualitative methods to generate nuanced results:

> They struggle too much to do what already they should be doing, which is SDR, KIIs and direct observation. If you add in there anything which is text, we do not recommend it, even now. . . . It's more important to be fast than to know the little nuances and differences between the gender and the old and the young and so on. (KI #23)

Finally, very short interviewer training sessions were cited by most respondents as a significant limitation for qualitative interview methods. KIs explained that, compared to quantitative assessment methods, qualitative methods require significantly more time to cover issues such as probing, empathy, and to ensure a thorough understanding of the issues being covered during the interview. According to KIs, several days should be devoted to conducting such training sessions—particularly because interviewers recruited during a crisis usually do not have previous experience conducting qualitative (or even quantitative) interviews. By contrast, respondents said that even a single day of training is not always possible due to budget or time constraints:

> There have been cases in the past where we would do a one-hour training, and that is the reality of NGOs sometimes. . . . So, a day is already an improvement on what was the case in some areas. Not saying it's right, I'm just speaking the truth. (KI #14)

Cluster 3.2: Notetaking/Transcript Challenges. Manually written notes of responses to OEQs are often of very poor quality. KIs said this was often the result of insufficient interviewer training, fatigue, as well as the challenges of writing down qualitative responses. Notes on OEQ responses are mostly written on paper, entered into structured Excel tables if using laptop computers, or entered via touchscreen keyboards into KoboToolbox or other CAPI forms. Respondents mentioned multiple specific challenges, most notably notes being too curt to be comprehensible, too verbose for quick analysis, or being hard to understand due to typos. One respondent noted that some interviewers often wrote "XXX" into text fields to be able to continue with the CAPI instrument. Such challenges can make qualitative notes hard to trust during analysis:

> Every time we tried it was a total disaster [laughs] it was a total failure. It requires so much. . . . Are they transcribing properly? Are they interpreting what the person is saying? They are doing their eighth interview today. They're just talking whatever. They're tired. So, it's never actually worked really. We tried that way. (KI #21)

In research settings, verbatim transcripts of qualitative interview responses are typically created post-interview, based on an audio recording. However, this practice is exceedingly rare during HNA, as noted one KI: "I don't see that. I have yet to see that in an operation that I've worked in" (KI #16).

Respondents cited the amount of time needed to create the transcript, creating a translation, and then condensing or coding it. Some also mentioned that the creation of audio files during interviews can be seen as adding a potential layer of risk to the respondent due to the lack of anonymity, preventing their wider use: "We often try not to record much just because

from experience, people are very nervous about recordings. . . . We guide our teams to only do that if they really know people are comfortable" (KI #14).

Cluster 3.3: Language Issues. A significant gap exists between many of the languages spoken by affected communities and those of the responding humanitarian organizations. This drastically hampers the ability to collect and analyze qualitative data. Respondents cited many examples, especially in African and Asian countries where foreign staff are usually unfamiliar with any of the local—or even national—languages spoken: "We had people that spoke only English and then we find one translator from English to Portuguese and another one from Portuguese to the local language" (KI #20).

Often, local staff stand in as impromptu translators, which may further reduce the quality of the output. Some KIs worried about the time this translation requires, but most expressed concern over the potential loss of meaning throughout the data collection cycle. Reflecting on a recent assessment in Mexico, one respondent said:

> How well they understood . . . the initial training with me, versus how well that person reads English; when I'm using a Haitian interviewer to speak to Haitians in French, but being trained in Spanish . . . I don't think it's at all perfect. (KI #04)

Due to the lack of translators and resources, written notes taken are often not translated into a language spoken by the person analyzing the data. As a result, they are often not considered during the analysis due to time constraints. In particular, this happens when interviewers need to write down the response after choosing "other" in a pre-coded OEQ:

> Most of the time, [the notes] don't get translated. . . . And by then, because they're so used to the systems where you can collect that on KoboToolbox, and then it goes straight into whatever automated system

> you've created, and you've got a graph, and you have your results already, they don't want to wait. Those steps slow the process down. And they also are extremely costly. (KI #05)

Even when qualitative data are analyzed, basing the analysis solely on translations may miss part of the story. Some respondents argued that in many countries, local staff are rarely involved in analysis in the source language even though they could better contextualize responses or clarify misunderstandings.

> Why don't we show the qualitative data so that it must be in the native language? . . . We must be able to do it in . . . Swahili, in Farsi, in Somali—not just in English, French, Arabic or Spanish. That's not good enough. There are so many words that change the entire meaning, or it will not exist in a certain vocabulary, that can completely screw your assessment. (KI #21)

Cluster 3.4: Systemic Challenges. Rather than build upon or complement each other, qualitative and quantitative methods are frequently not integrated in a mixed-methods approach. Respondents noted that when qualitative interviews are used, they are often done in lieu of a household survey, or vice versa, to maximize resources:

> We wind up shoving more of certain sectors into the qualitative, and some of the others out into the quants data. So, we can ask most economic-type questions in quant data. And so, most of that gets sidelined and pushed out of the qualitative data, which creates problems sometimes. (KI #04)

Many organizations struggle with biases against qualitative methods. Many respondents pointed out that these biases can be found at all levels, ranging from assessment specialists to senior leadership, sometimes pervading an entire organization's "culture" around data collection.

KIs said there is a widespread belief among many humanitarian decisionmakers that only household data (preferably from large samples) and numeric graphs can be used for operational decision-making. As a result, qualitative results were said to be excluded from reports, or not thoroughly analyzed in the first place. Worrying findings can also be brushed aside if they stem from qualitative sources. One KI paraphrased the response received from a large agency:

> Just because 12 households or whatever told you [about deficiencies in the quality of food assistance] that doesn't really mean anything in a response of one million people. Our teams who work for us, the organization providing the food, found that there wasn't a problem. So, it's fine. (KI #02)

There has been a drive to conduct HNA increasingly remotely by using computer-assisted telephone interviewing (CATI) (Gibson et al., 2017), a trend accelerated in 2020 and 2021 due to the COVID-19 pandemic. Several KIs noted that many staff were initially enthusiastic about no longer needing to travel, as noted KI #22: "We don't need to travel anymore, we can do everything remotely."

Others noted that increasing remote data collection may become the norm, and that survey respondents may provide less accurate information to maximize assistance. Phone interviews needed to be shorter than face-to-face interactions, according to respondents, as interviewees are considered to be unwilling to participate if the call is too long. This has reduced the amount of qualitative data collected in many cases, as most or all OEQs were removed.

Similarly, lack of field site access, which has been compounded by the COVID-19 pandemic, has led to a push for more automation and secondary data analysis. Respondents said there was an unrealistic expectation by some donors and organizations to focus on automating the analysis of social media posts (through the Data Entry and Exploration Platform (DEEP)

(DEEP, 2021)), which they said came at the expense of collecting more qualitative primary data. Some KIs also pointed out that mistakes due to automated coding of qualitative data could lead to an overall lack of trust in the results or—possibly worse—acceptance of results without proper human verification.

> If the algorithm is so complex, I don't really know how I come to this conclusion as a big likelihood that you're not going to take any decisions based on that. Particularly in the humanitarian sphere. So somehow, people need to be involved also in those ambitions. And obviously, there's the hope that a lot of it is being automatized. (KI #06)

4.2.6. Thematic Group 4: Challenges Specific to Pre-coded Open-Ended Questions (OEQs)

Pre-coded OEQs straddle the divide between quantitative and qualitative methods. In structured surveys, pre-coded OEQ can be understood as qualitative questions that require the interviewer to code the response on the spot into one or more pre-defined choices in the questionnaire instrument. Because of this immediate coding practice, responses to OEQ are immediately converted into quantitative data during the survey. Pre-coded OEQs often also have an option to enter text for responses not reflected among the questionnaire choices. Such text, typically entered after choosing "other" among the pre-coded responses, then requires qualitative analysis methods, which may in turn result in additional codes that are then added to the original question options during the reporting stage. Less commonly, structured surveys may also have OEQ that only allow entering a few words to summarize the response. This study considers the themes related to pre-coded OEQ as a separate component as they highlight issues that KIs often discussed separately from quantitative and qualitative ones. Four challenges unique to coding OEQs are discussed below.

Many respondents discussed the challenge of pre-coding the correct responses to OEQs into survey instruments. Respondents identified the need to establish response options during initial research, such as conducting focus group discussions, secondary data review, piloting, or conversations with local staff as validation methods. However, several respondents noted that this is often neglected in practice. Without establishing response options, several KIs pointed to the risk of biased results being considered as evidence:

> So, the ideal needs assessment [at the UN level] is a household survey. And it's quantitative. But they're collecting qualitative data in a quantitative format. And we make these ridiculous assumptions that we know the answer options we should be giving without testing them, and we roll them out globally and keep using the same ones. And because we use them in so many places, we consider that evidence that they work. But we never often go in and actually check that that's what we should be asking. So, there's problems with how we phrase them. There's bias built into the questions we use, that we repeatedly implement. (KI #05)

Compounding the problem of excluding important options in pre-coded OEQs, interviewers often select the wrong item, either because of personal biases, fatigue, or lack of familiarity with the terminology. KIs said that interviewers are sometimes not sufficiently trained to understand the intention behind each pre-coded option in the instrument, too often choosing "other" rather than considering which pre-coded response fits best.

> OEQs were often considered by KIs to be an obstacle to collecting data from larger samples. This is due to the additional time needed for respondents to answer such questions "because we're trying to get larger numbers of responses and to get a better picture of more people" (KI #03).

One organization had even eliminated all OEQs from some of their MSNAs to save time, as surveys needed to be conducted over the phone due to the COVID-19 pandemic. Others said that to save time, they removed the option for interviewers to enter text after the "other" option was selected—thereby converting OEQs to closed-ended questions.

Finally, many respondents admitted that text responses entered for OEQs were not systematically analyzed, or not analyzed at all, for a large proportion of HNAs. For example, KI #04 stated: "You mean, translate and fully analyze it? Maybe 10% of the time."

Many respondents underlined the importance of collecting open-ended data and conducting proper analysis for methodological completeness. However, most KI reported that that manual analysis (and often translation) of these notes is either not done well or not done at all, "because you focus on your 90% of the data that's easily processed" (KI #09).

4.2.7. Thematic Group 5: Suggestions for Improving the Role of Qualitative Interview Methods in HNA

Given the many challenges highlighted by KIs, many respondents also proposed several concrete solutions during interviews. Eight themes were identified that cover various operational, methodological, operational, and technical ideas.

Most respondents proposed that interviewers be given more training, ideally over three days, to be able to collect better (or any) qualitative data. Funding was universally identified as the limiting factor in conducting longer training sessions.

> You need much more rigorous training. Right now, trying to get three days of training for enumeration is a stretch; for qualitative data, you would need way more than that. And it's just that timeframe. And funding is a major problem. (KI #05)

KIs emphasized that involving national staff in the design and analysis stages would also be essential to ensure meaning is preserved and not lost by one or more rounds of translation. Relatedly, many KIs suggested that humanitarian organizations should be hiring additional staff or training existing teams so that more HNAs are led by people who are also trained in qualitative methods. In some organizations, this would require a significant change from the current focus on recruiting more analysts with quantitative skills, particularly Information Management Officers (IMOs):

> We have a massive problem with assessments because we're not hiring the right people to do them. So, we're focusing on that IMO type, who don't have the right background. . . . Even if you want people with serious humanitarian background, hire your M&E people to run the assessments. . . . Make sure that you've got some people who know how to design it and train people on qualitative data collection, because it's a lot harder to do. (KI #05)

Several respondents suggested that many household surveys should have fewer questions, for ethical and resource reasons. One respondent argued that, in their experience, too much quantitative data were collected but never used, and estimated that "we probably analyze 20% of our data properly" (KI #08). Some were highly critical with the increasing amount of data collected as part of annual MSNAs across all long-term crises, arguing instead that qualitative methods should be used to collect deeper information:

> For protracted crises, I do not understand why we are putting so many questions in our questionnaires. . . . It's insane. 450 variables. 900 variables. It's crazy. It's hours of discussion. It's bullshit. Everything that you ask after 45 minutes, it's bullshit. The guy just wants to leave or to be somewhere else. . . . I think it's not smart. It's simply not smart what

> we are doing right now. Putting [US$] 1.5 million into a field assessment, there is barely a secondary data review done beforehand. There are barely any qualitative interviews being done in advance to help shape the questionnaire. . . . For me, it's a real waste of money and I find it damaging, harmful, non-ethical, irresponsible. . . . In the field, we should have much more qualitative data. (KI #23)

Several respondents imagined that both household surveys and KIIs could integrate a much larger proportion of OEQs to collect relevant information more efficiently. They argued that combining quantitative and qualitative approaches in surveys would generate data that are at once timely, representative, and provide more depth. Some KIs proposed creating short audio recordings for each OEQ alongside quantitative questions in CAPI instruments. One respondent with significant experience conducting MSNAs suggested that such data would be more useful even while allowing them to remove many quantitative questions:

> What if I can have a shorter questionnaire, and then I ask . . . "Which are your three priorities?" And the person says "food." And then you tell them "Okay, great. Can you explain me why?" I could probably cut my survey in half. . . . I could get rid of at least 20–30 questions, easily. (KI #21)

Respondents also discussed a number of technical proposals. Several KIs argued that having audio recordings of interview responses would be beneficial for multiple purposes. This includes the ability to conduct quality control of interviews, derive more data from a response, keep the interview shorter, and to create transcripts. One respondent said, "If I could choose, I would always go for transcripts and recordings" (KI #13) but added that technical and resource constraints prevented them from doing this more widely.

Without specific prompting, several KIs also proposed that HNA should make use of automated transcription, translation, and analysis tools as potential solutions for collecting more and better qualitative data. Such tools use natural language processing (NLP), a set of methodologies using artificial intelligence (AI), which include automated speech recognition (ASR), machine translation (MT), as well as different types of content analysis (Kreutzer et al., 2020). The potential use of ASR for transcribing interview responses was mentioned by several respondents as a way to increase the number of qualitative interviews and to overcome the poor quality of manually written interview notes. Some KIs argued that ASR would reduce the amount of time spent by local teams who are only able to transcribe a small number of audio files. Others mentioned that ASR-generated transcripts would be superior to those created by local teams because the "human bias that comes in, whether we like it or not, whoever is listening to the recording may not choose to transcribe something because they don't think it's important" (KI #09).

Several KIs noted that automatically translating interview transcripts or manually entered notes to OEQs would be crucial for bridging language gaps to analyze more qualitative data to inform HNAs. Several respondents cited the idea of employing MT to ensure translations are available more quickly:

> If you had Google Translate, and you could export the qualitative wording . . . We would copy and paste that into our Excel file and start doing our analysis that way around. . . . But I think the biggest hurdle is the translation. (KI #14)

Two respondents cited concerns related to ASR and MT due to the lack of support for certain languages spoken in many humanitarian crises, saying that investments for creating models for these languages are crucial. One respondent argued that tools and processes are

needed to ensure that human translators can do manual transcription and translation while also using this work to improve language models.

Several respondents proposed the use of "automated" methods for analyzing qualitative data, referring to various NLP techniques. In their view, use of such tools for HNA "would be the big lightbulb, the big key that would change everything" (KI #02). One KI argued that better tools for automatically coding responses in qualitative interviews would allow collecting data from much larger samples:

> If we had a way to take qualitative information and analyze it to the depth of what we can do with quantitative information, we could interview 10 times as many people as we currently do. And we could then have far more representative data about a population at a much more intimate and deeper level. (KI #03)

One respondent with significant qualitative experience explained that while current technology does not yet match in-depth human coding, it could augment what a human analyst is able to do in a short time frame:

> AI does give you those options. . . . If it was a choice between me manually coding it [over several weeks], and then just me skimming over it and getting the general ideas and writing report, then the auto code in NVivo would be better than that. (KI #10)

4.3. Discussion

The aim of this study was to explore current approaches to conducting HNA, constraints related to different interview methods—with particular attention to qualitative methods, and solutions for improving the use of qualitative interview methods in HNAs. In-depth interviews

with 23 purposively sampled KIs from a wide array of humanitarian organizations showed widespread support to the idea that qualitative information should be used more widely to complement quantitative data in order to accurately assess population needs in humanitarian crises. The information obtained from interviews identifies many of the significant obstacles to collecting qualitative data due to extremely challenging operational circumstances, institutional biases, and limited resources. As a result, most in-depth assessments, such as MSNA, use overwhelmingly quantitative methods to measure information that may be better and more fully captured using qualitative methods, leading to severe limitations on the reliability and accuracy of such data. Even when qualitative data is collected, KIs explained that even though much more data is collected for HNA than are ultimately used, qualitative data is disproportionately under-used compared to quantitative data.

These findings may lead some to conclude, perhaps cynically, that qualitative methods should be abandoned altogether in humanitarian needs assessments. However, most KIs underlined the importance of qualitative data that cannot be replaced by quantitative data alone. Respondents pointed to several suggestions. Four concrete solutions emerged from the interviews with KIs which, if scaled correctly, could significantly improve the overall humanitarian response:

1. Acknowledge limitations of quantitative data
2. Increase number of staff with qualitative skills
3. Methodological innovations
4. Technological innovations

4.3.1. Acknowledge Limitations of Quantitative Data

First, KIs pointed to many systemic biases against qualitative and in favor of quantitative data that need to be addressed openly. Each approach has unique advantages and shortcomings, whereas mixed-methods approaches that leverage the strengths from different methodological approaches would provide the strongest HNA evidence to inform humanitarian assistance. Quantitative approaches to measuring complex qualitative issues (including "What are your three priority needs?") can result in precise figures and graphs (ACAPS, 2012). However, precise often does not equal accurate (ACAPS, 2013). Humanitarian organizations should acknowledge that many complex issues cannot be measured reliably with current approaches. KIs pointed out that asking rapid-fire, closed-ended questions risks offending affected people and may negatively influence their responses. They also pointed out that much more quantitative data is collected than is utilized to inform operations. Qualitative approaches, such as patiently listening to the responses to OEQs, can convey empathy and dignity, and can also lead to better data.

This study shows that, similar to quantitative data, a large amount of qualitative data collected during HNAs are not analyzed or included in reports. KIs ascribed this to severe time and capacity limits, but also reported receiving a strong signal from leadership levels that it is not useful in reports. Instead, we argue that organization leadership and donors should insist that existing qualitative data be analyzed and used more prominently in reports to complement, explain, or show gaps in quantitative findings. Likewise, donors should insist that more qualitative data are collected and mobilize financial and staffing resources to do so. KIs pointed to donors' willingness to expend significant resources for quantitative MSNA, which can cost in excess of US$ 1 million, suggesting that many of the financial requirements could be met through reallocation rather than additional funding.

KIs argued that annual MSNAs, which typically collect hundreds of indicators for all humanitarian sectors, are often not useful for decision-making in actual field conditions. Their primary purpose is to inform each country's annual Humanitarian Response Plan (HRP) (Humanitarian Programme Cycle, 2021), though this uses only a fraction of the collected data. Several KIs were critical of the recently created Joint Intersectoral Analysis Framework (JIAF), which aims to make MSNA data more comparable across crises. However, respondents with personal involvement argued that initiatives like the JIAF overemphasize quantitative data for the purpose of allowing comparisons of severity between very different crises. The emphasis on MSNAs and the creation of the JIAF are the results of the 2016 Grand Bargain commitment to "improve joint and impartial needs assessments," which explicitly focused on improving the HRP (IASC, 2019b). Despite the significant amounts of energy and resources invested in these processes, it is important to acknowledge that many operational decisions are still constrained by information gaps that cannot be addressed through ever larger and more detailed MSNAs. As the data in this study show, operational decisions require faster data collection with contextualized instruments, including more qualitative data. Ongoing work to improve HNA under the Grand Bargain process, which to date has focused heavily on quantitative tools, needs to address its blind spots regarding qualitative data. This could include adopting existing guidance documents on how and when qualitative methods should be used, such as the recently-created Counter-trafficking in Emergencies guide by IOM or the Global Protection Cluster's Protection Analytical Framework training materials (Global Protection Cluster, 2021; IOM, 2021).

Humanitarian organizations should also better integrate teams responsible for community engagement and accountability into HNA workstreams. The qualitative approaches used by these

units to understand population feedback should become a core part of HNAs—not implemented only after programming has already begun, or performed merely in parallel to assessments.

4.3.2. Increase Number of Staff with Qualitative Skills

Interviews conducted for this study have uncovered a lack of qualitative capabilities as a major hurdle, including poor interviewing and notetaking skills, insufficient capacity, and improper analysis of qualitative data. KIs frequently pointed to hard budgetary constraints on training of interviewers. Instead, training sessions should be expanded for better assimilation of qualitative ways of asking questions, probing, and taking detailed and reliable notes. These skills also apply to most household surveys, which can contain many open-ended (and therefore qualitative) questions. To expand analytic capacity, more staff with qualitative skills should be recruited, and training should be prioritized for those involved in HNA. We believe humanitarian organizations should advocate for dedicated funding to ensure that training sessions are expanded whenever possible.

KIs recommended that HNAs not be led by purely quantitative staff in the "information management" (IM) role. In many organizations, "IM expert" has become synonymous with "assessment expert," despite the usually very different roles and qualifications (IASC, 2019a). Instead, respondents suggested better training of existing staff (including IM specialists), recruiting of new staff with qualitative training, and reallocating HNA to monitoring and evaluation teams with experience in mixed methods.

Many KIs reported that staff with qualitative skills often cannot speak the local languages (requiring everything to be translated into English), and that the staff with local language and cultural knowledge often do not have qualitative skills. In fact, one of the primary drivers of the enthusiastic adoption of quantitative methods for measurement of qualitative information has

been their ability to circumvent the language barriers between the international team, local staff, and respondents. Conducting more training of staff (particularly local) in qualitative analysis methods and prioritizing of trained translators' involvement throughout the HNA process are the best means of closing this gap. However, if translations are needed for international staff, organizations should plan—and budget—for the use of professional translators whenever possible to guard against misunderstandings and biases. Today, too many translations "are done by just that guy who sits there and knows that language" (KI #05), which limits both the speed and most importantly, the reliability of qualitative work.

4.3.3. Methodological Innovations

More key informant interviews should be conducted qualitatively using semi-structured instruments, especially outside the immediate circumstances of sudden-onset emergencies. Current humanitarian methods training programs and guidelines widely consider KIIs as a qualitative method whose purpose is to answer deeper questions. For example, UNHCR's standard-setting Needs Assessment Handbook lists KII only as a qualitative research method (UNHCR, 2017, p. 111). In practice, however, respondents in this study showed that most KIIs in HNA are, in fact, increasingly conducted quantitatively using structured questionnaires. Time pressures associated with humanitarian crises have, over time, led some to recommend to "keep [KIIs] focused and limit the number of questions asked" (ACAPS, 2011, p. 14) to rapidly obtain quantitative baseline information. Methods training and reports should clearly differentiate between structured quantitative KIIs and semi-structured qualitative KIIs, and their respective strengths and weaknesses.

Taking this idea to the next level, several KIs proposed to adopt a new mixed-method approach for collecting qualitative and quantitative data by integrating more uncoded OEQs into

structured data collection instruments (such as household surveys or KII). Such a hybrid approach would reduce the number of pre-coded or closed-ended questions and instead require analyzing responses to a subset of questions based on transcripts rather than interviewers choosing pre-coded options alone. This type of mixed-methods survey approach would require both quantitative and qualitative analysis of the different indicators in the same dataset. Implementing such surveys may require certain technological innovations for collection and processing of large numbers of audio files (see below).

Previous research has shown that using pre-coded options in structured survey questions can lead to different results compared to doing content analysis later on (Ozuru et al., 2013), whereas linguistic and conceptual challenges can also result in significant mistakes by enumerators (Translators Without Borders, 2018a). We argue that, in practice, this new mixed-methods hybrid survey method should include four specific steps:

1. Household or KI interview instruments are designed to include more OEQs. Where possible, the number of closed-ended questions is reduced.
2. Interviewers record the audio to OEQs while also entering quantitative answers to other questions in the survey instrument, and both are stored together using CAPI tools.
3. Native speakers, supported by NLP, create rapid transcripts each day (and translations, if needed), which are saved as part of the survey dataset.
4. Assessment staff assign codes to transcripts, which become part of the survey dataset. Ideally, this step would be performed by local staff in the interview language to avoid challenges with translating original responses.

Where audio recordings are not possible, this Step 2 may instead require the interviewer to take detailed notes. If typing on small mobile devices in CAPI questionnaires is a hindrance for speed or quality reasons, handwritten notes could be taken that are then entered during Step 3. Two KIs proposed that recording, transcribing, and analyzing audio could be required for only a subset of respondents to make the additional work more manageable. Such sub-sampling could be done at random or through conditional display of questions based on the specific profile of a respondent (e.g., female or child-headed households). Pilot assessments should be conducted to develop a stable methodology, to document the time and resources required, and to establish best practices that can be used in training.

As recent work to rapidly analyze qualitative feedback related to the Ebola response in DR Congo showed, even when qualitative analysis is done rapidly, it still requires significant resources for manual processing of data (Earle-Richardson et al., 2021). We note that rapid qualitative analysis—a group of alternative methods for analyzing responses to qualitative data more quickly (Vindrola-Padros & Johnson, 2020; Watkins, 2017)—may also have the potential of achieving more robust datasets at reduced costs and in less time compared to traditional qualitative analysis (Taylor et al., 2018).

4.3.4. Technological Innovations

Technological innovations will be essential to support the other recommendations mentioned above, particularly the proposed mixed-methods survey approach. Without them, major qualitative undertakings to understand populations will remain rare, slow, and expensive. Data from this study point to four specific technological innovations, as proposed by KIs:

1. CAPI or CATI tools should be used for easy and secure recording of audio for any number of OEQs. Ideally, this should be done at the same time as entry of

quantitative data. Manual notetaking needs to remain available if field contexts dictate it. To mitigate ethical risks, audio files should be stored no longer than necessary.

2. Whenever possible, ASR can be used to speed up the transcription process. Manual correction by native speakers will be important to ensure against errors and to guard against biases (e.g., people who speak softly, have strong accents, etc.).
3. If needed, MT should be employed to assist in the translation of open-ended responses. As with ASR, manual verification is essential to ensure accuracy. For languages for which ASR or MT are unavailable, manual transcription or translation should be made easy. Investments should be made to expand ASR and MT models to cover more of the languages spoken in humanitarian crises and to create an alternative option for processing of sensitive data without the use of commercial services.
4. NLP techniques for content analysis need to become more accessible to humanitarian users so that relevant information from qualitative responses may be classified more rapidly.

The field of NLP is advancing rapidly, making methods for content analysis (such as classification, information extraction, and topic modeling) more widely available to specialized researchers (Cunningham et al., 2013; Loper & Bird, 2002; McCallum, 2002; Mihalcea, 2004). Such approaches are increasingly being used in health and social science research (Roberts et al., 2021) and can be particularly useful to support human coding of OEQs (Popping, 2015). Several KIs suggested that content analysis methods supported by NLP should be explored further,

though some were concerned about the potential ethical risks. More research is needed to pilot such methods and establish best practices for verifying the accuracy of AI-assisted analysis methods.

The above suggestions, based on data obtained from KIs in this study, can be implemented incrementally and independently of each other. They have the potential to improve the quality of HNAs and therefore of making humanitarian assistance more relevant, timely, and adapted to local requirements. More broadly, these suggestions stand to benefit survey methods throughout humanitarian response, global health, and related fields of research, policy and practice. Finally, technological innovations, particularly using NLP and other forms of AI, should be created to mitigate a growing list of ethical implications related to processing data from affected people (Kreutzer, 2021).

Strengths of this research include a diverse sample of KIs with deep experience in HNA from a large geographic area and wide sampling across an array of organizations providing humanitarian assistance. The results fill an important gap in the academic literature where empirical data about practices and challenges in HNA are lacking. Further, the findings align well as a significant contribution to the evidence base need to support the commitments articulated in the Grand Bargain to improve needs assessments and to include affected people in decision-making (WHS, 2016). Relevant limitations include relative over-sampling of KIs from large international humanitarian organizations, as well as under-sampling of KIs from Africa, East Asia, and Latin America. Although KIs represented 17 organizations, it cannot be assumed that the data reflect all views within these large institutions, nor that perspectives are representative of other organizations.

4.4. Conclusion

This study found that despite a rapidly growing volume of quantitative data, the gap in understanding between the people affected by humanitarian crises and those charged with providing assistance has been widening in recent years—particularly since the onset of the COVID-19 pandemic. The role of qualitative methods in humanitarian needs assessment remains limited compared to the role of quantitative methods, due to a combination of systemic, technological, and methodological factors. Expanding the use of qualitative data to inform humanitarian action has strong support across the humanitarian sector. Global guidance, informed by cross-sectoral consultation, is urgently needed to improve the increasingly challenging task of understanding population needs. The authors' proposals derived from this study provide a roadmap for strengthening the collection of qualitative interview data throughout the humanitarian program cycle in humanitarian crises.

4.5. References

ACAPS. (2011). *Technical brief: Direct observation and key informant interview techniques for primary data collection during rapid assessments.* ACAPS. https://www.acaps.org/sites/acaps/files/resources/files/direct_observation_and_key_informant_interview_techniques_for_primary_data_collection_during_rapid_assessments_october_2011.pdf

ACAPS. (2012). *Qualitative and quantitative research techniques for humanitarian needs assessment*. ACAPS. https://www.acaps.org/sites/acaps/files/resources/files/qualitative_and_quantitative_research_techniques_for_humanitarian_needs_assessment-an_introductory_brief_may_2012.pdf

ACAPS. (2013). *How sure are you ?* ACAPS. https://www.acaps.org/sites/acaps/files/resources/files/how_sure_are_you-judging_quality_and_usability_of_data_collected_during_rapid_needs_assessments_august_2013.pdf

ACAPS. (2014). Humanitarian Needs Assessment: The Good Enough Guide. In *Geneva: ACAPS*. ACAPS. https://doi.org/10.3362/9781780449357

ACAPS. (2019). *Lessons Learned: Needs Assessments in Cox's Bazar*. ACAPS. https://www.acaps.org/sites/acaps/files/products/files/20190404_acaps_npm_coxs_bazar_analysis_hub_needs_assessments_lessons_learned.pdf

Al-Amer, R., Ramjan, L., Glew, P., Darwish, M., & Salamonson, Y. (2016). Language translation challenges with Arabic speakers participating in qualitative research studies. *International Journal of Nursing Studies*, *54*, 150–157. https://doi.org/10.1016/j.ijnurstu.2015.04.010

Banatvala, N. (2000). Conflict and health: Public health and humanitarian interventions: developing the evidence base. *BMJ*, *321*(7253), 101–105. https://doi.org/10.1136/bmj.321.7253.101

Belgrave, L. L., & Seide, K. (2018). Grounded Theory Methodology: Principles and Practices. In P. Liamputtong (Ed.), *Handbook of Research Methods in Health Social Sciences* (pp. 1–18). Springer, Singapore. https://doi.org/10.1007/978-981-10-2779-6_84-1

Bowden, A., & Fox-Rushby, J. A. (2003). A systematic and critical review of the process of translation and adaptation of generic health-related quality of life measures in Africa, Asia, Eastern Europe, the Middle East, South America. *Social Science & Medicine*, *57*, 1289–1306. https://doi.org/10.1016/S0277-9536(02)00503-8

Building Markets, & Orange Door Research. (2018). *What is the Point if Nothing Changes: Current Practices and Future Opportunities to Improve Remote Monitoring and Evaluation in Syria*. New York, NY: Building Markets.

Chataigner, P. (2017). *Analytical Framework Review report- A framework analysis of analytical frameworks*. OCHA. https://assessments.hpc.tools/km/analytical-framework-review-report

Chun Tie, Y., Birks, M., & Francis, K. (2019). Grounded theory research: A design framework for novice researchers. *SAGE Open Medicine*, *7*, 205031211882292. https://doi.org/10.1177/2050312118822927

Corbin, J., & Strauss, A. (2015). *Basics of Qualitative Research* (4th editio). Sage Publications, Inc.

Cosgrave, J. (2009). Humanitarian funding and needs assessment. In *Humanitarian Response Index 2008: Donor Accountability in Humanitarian Action* (pp. 75–88). https://doi.org/10.1057/9780230584617

Creswell, J. W. (2014). *Research Design* (4th ed.). Sage.

Cunningham, H., Tablan, V., Roberts, A., & Bontcheva, K. (2013). Getting More Out of Biomedical Documents with GATE's Full Lifecycle Open Source Text Analytics. *PLoS Computational Biology*, *9*(2), e1002854. https://doi.org/10.1371/journal.pcbi.1002854

DEEP. (2021). *Data Entry and Exploration Platform*.

Deutscher, G. (2010). *Through the Language Glass: Why the World Looks Different in Other Languages*. Henry Hold and Company.

Earle-Richardson, G., Erlach, E., Walz, V., Baggio, O., Kurnit, M., Camara, C. A., Craig, C., Dios, L. R., Yee, D., Soke, G. N., Voahary, I., & Prue, C. E. (2021). New mixed methods approach for monitoring community perceptions of ebola and response efforts in the democratic Republic of the Congo. *Global Health Science and Practice*, *9*(2), 332–343. https://doi.org/10.9745/GHSP-D-21-00144

Gerdin, M., Chataigner, P., Tax, L., Kubai, A., & von Schreeb, J. (2014). Does need matter? Needs assessments and decision-making among major humanitarian health agencies. *Disasters*, *38*(3), 451–464. https://doi.org/10.1111/disa.12065

Gibson, D. G., Pereira, A., Farrenkopf, B. A., Labrique, A. B., Pariyo, G. W., & Hyder, A. A. (2017). Mobile Phone Surveys for Collecting Population-Level Estimates in Low- and Middle-Income Countries: A Literature Review. *Journal of Medical Internet Research*, *19*(5), e139. https://doi.org/10.2196/jmir.7428

Global Health Cluster. (2017). *Operational Guidance on Accountability to Affected Populations (AAP)*. https://www.humanitarianresponse.info/system/files/documents/files/cluster_coordination_reference_module_2015_final.pdf

Global Protection Cluster. (2021, August 11). *Protection Analytical Framework*. https://www.globalprotectioncluster.org/2021/08/11/protection-analytical-framework/

GPPI, & INSPIRE Consortium. (2019). *Methodology to Assess Coordinated Multi-Sector Needs Assessments*. Global Public Policy Institute. https://www.gppi.net/project/assessing-needs-assessments

Gray, L. M., Wong-Wylie, G., Rempel, G. R., & Cook, K. (2020). Expanding qualitative research interviewing strategies: Zoom video communications. *Qualitative Report*, *25*(5), 1292–1301. https://doi.org/10.46743/2160-3715/2020.4212

Humanitarian Programme Cycle. (2021). *Humanitarian Response Plan Guidance*. OCHA.

IASC. (1999). *Protection of Internally Displaced Persons*. https://interagencystandingcommittee.org/system/files/legacy_files/protection_of_internally_displaced_persons_inter_agency_standing_committee_policy_paper_0.pdf

IASC. (2015). Multi-Sector Initial Rapid Assessment Guidance. In *Geneva: Inter-Agency Standing Committee*. https://interagencystandingcommittee.org/system/files/mira_2015_final.pdf

IASC. (2019a). *Ensuring Data and Analysis is Useful and Usable for Response - Tools*. Inter-Agency Standing Committee. https://interagencystandingcommittee.org/system/files/ensuring_data_and_analysis_is_useful_and_usable_for_response_-_tools.pdf

IASC. (2019b). *Workstream 5: Tools and guidance for advancing coordinated needs assessment and analysis through the Grand Bargain*. Inter-Agency Standing Committee. https://interagencystandingcommittee.org/improve-joint-and-impartial-needs-assessments/workstream-5-tools-and-guidance-advancing-coordinated

ICRC. (2016). *Protracted conflict and humanitarian action | ICRC*. International Committee of the Red Cross. https://www.icrc.org/en/document/protracted-conflict-and-humanitarian-action

International Committee of the Red Cross. (2019). *Accountability to Affected People Institutional Framework*. https://www.icrc.org/en/publication/accountability-affected-people-institutional-framework

International Office for Migration. (2021, November 23). *Displacement Tracking Matrix*. https://dtm.iom.int/

IOM. (2021). *Counter-trafficking in Emergencies: Information Management Guide*. International Officer for Migration. https://publications.iom.int/books/countertrafficking-emergencies-information-management-guide

Kobo. (2022). *KoboToolbox* (2.021.45). https://www.kobotoolbox.org

Kreutzer, T. (2021). *Ethical Implications Related to Processing of Personal Data in Humanitarian Crises: A Scoping Review*.

Kreutzer, T., Vinck, P., Pham, P., An, A., Appel, L., Deluca, E., Tang, G., Alzghool, M., Hachhethu, K., Morris, B., Crowley, J., & Orbinski, J. (2020). Improving Humanitarian Needs Assessments through Natural Language Processing. *IBM Journal of Research and Development*, *64*(1). https://doi.org/10.1147/JRD.2019.2947014

Loper, E., & Bird, S. (2002, May 17). NLTK: The Natural Language Toolkit. *Proceedings of the ACL Workshop on Effective Tools and Methodologies for Teaching Natural Language Processing and Computational Linguistics*. http://arxiv.org/abs/cs/0205028

MAXQDA 2020. (2019). VERBI Software. https://www.maxqda.com

McCallum, A. K. (2002). *MALLET: A Machine Learning for Language Toolkit*.

http://mallet.cs.umass.edu/

Mihalcea, R. (2004). Graph-based ranking algorithms for sentence extraction, applied to text summarization. *Proceedings of the ACL 2004 on Interactive Poster and Demonstration Sessions*. https://doi.org/10.3115/1219044.1219064

Morrow, N., Mock, N., Bauer, J. M., & Browning, J. (2016). Knowing just in time: Use cases for mobile surveys in the humanitarian world. *Procedia Engineering*, *159*(June), 210–216. https://doi.org/10.1016/j.proeng.2016.08.163

O'Brien, B. C., Harris, I. B., Beckman, T. J., Reed, D. A., & Cook, D. A. (2014). Standards for reporting qualitative research: A synthesis of recommendations. *Academic Medicine*, *89*(9), 1245–1251. https://doi.org/10.1097/ACM.0000000000000388

OCHA. (2021a). *Global Humanitarian Overview 2022*. Office for the Coordination of Humanitarian Affairs. https://doi.org/10.18356/9789210012423

OCHA. (2021b). *Joint intersectoral analysis framework 1.1*. UNOCHA. https://doi.org/10.18356/9789214030751c020

Okular Analytics. (2021). *Technical Review of 2020 MSNAs in Protracted Crises*. IMPACT Initiatives. https://www.reachresourcecentre.info/wp-content/uploads/2021/07/Okular_Report_MSNAs2020ExternalReview_July2021.pdf

Ozuru, Y., Briner, S., Kurby, C. A., & McNamara, D. S. (2013). Comparing comprehension measured by multiple-choice and open-ended questions. *Canadian Journal of Experimental Psychology*, *67*(3), 215–227. https://doi.org/10.1037/a0032918

Popping, R. (2015). Analyzing Open-ended Questions by Means of Text Analysis Procedures. *BMS Bulletin of Sociological Methodology/ Bulletin de Methodologie Sociologique*, *128*(1), 23–39. https://doi.org/10.1177/0759106315597389

Redmond, A. D. (2005). Needs assessment of humanitarian crises. *Bmj*, *330*(7503), 1320–1322. https://doi.org/10.1136/bmj.330.7503.1320

Roberts, C. h., Brindle, H., Rogers, N. T., Eggo, R. M., Enria, L., & Lees, S. (2021). Vaccine Confidence and Hesitancy at the Start of COVID-19 Vaccine Deployment in the UK: An Embedded Mixed-Methods Study. *Frontiers in Public Health*, *9*(November). https://doi.org/10.3389/fpubh.2021.745630

Spiegel, P. B., Burkle, F. M., Dey, C. C., & Salama, P. (2001). Developing public health indicators in complex emergency response. *Prehospital and Disaster Medicine*, *16*(4), 281–285. https://doi.org/10.1017/S1049023X00043430

Taylor, B., Henshall, C., Kenyon, S., Litchfield, I., & Greenfield, S. (2018). Can rapid approaches to qualitative analysis deliver timely, valid findings to clinical leaders? A mixed methods study comparing rapid and thematic analysis. *BMJ Open*, *8*(10), 5–7. https://doi.org/10.1136/bmjopen-2017-019993

Translators Without Borders. (2018a). *The words between us: How well do enumerators understand the terminology used in humanitarian surveys? A study from northeast Nigeria*. https://translatorswithoutborders.org/wp-content/uploads/2018/12/TWB_Nigeria_EnumeratorComprehension_Nov2018-1.pdf

Translators Without Borders. (2018b, April 19). *Communications Dashboard: Internally Displaced People in Northeast Nigeria*. https://translatorswithoutborders.org/communications-dashboard-internally-displaced-people-in-north-east-nigeria/

Translators Without Borders. (2019). *Language mapping: Putting communication needs on the map*. https://translatorswithoutborders.org/language-mapping/

UNHCR. (2006). *Master Glossary of Terms Rev. 1 United Nations High Commissioner for Refugees*. Geneva: UN High Commissioner for Refugees (UNHCR). http://www.refworld.org/docid/42ce7d444.html

UNHCR. (2017). *Needs Assessment Handbook*. http://needsassessment.unhcr.org/media/downloads/2017-05_NAH_BOOK_NEW.pdf

Vindrola-Padros, C., & Johnson, G. A. (2020). Rapid Techniques in Qualitative Research: A Critical Review of the Literature. *Qualitative Health Research*, *30*(10), 1596–1604. https://doi.org/10.1177/1049732320921835

Watkins, D. C. (2017). Rapid and Rigorous Qualitative Data Analysis: The "RADaR" Technique for Applied Research. *International Journal of Qualitative Methods*, *16*(1), 1–9. https://doi.org/10.1177/1609406917712131

WHS. (2016). *The Grand Bargain – A Shared Commitment to Better Serve People in Need*. Istanbul: World Humanitarian Summit. http://reliefweb.int/sites/reliefweb.int/files/resources/Grand_Bargain_final_22_May_FINAL-2.pdf

World Bank. (2020). *World Bank Country and Lending Groups*. The World Bank. https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups

4.6. Chapter 4 - Appendix 1

Standards for Reporting Qualitative Research Checklist

Based on O'Brien, B. C., Harris, I. B., Beckman, T. J., Reed, D. A., & Cook, D. A. (2014). Standards for reporting qualitative research: A synthesis of recommendations. *Academic Medicine*, *89*(9), 1245–1251. https://doi.org/10.1097/ACM.0000000000000388

#	Reporting item	Page number
	Title and abstract	
1	Title - Concise description of the nature and topic of the study Identifying the study as qualitative or indicating the approach (e.g., ethnography, grounded theory) or data collection methods (e.g., interview, focus group) is recommended	126
2	Abstract - Summary of key elements of the study using the abstract format of the intended publication; typically includes background, purpose, methods, results, and conclusions	126
	Introduction	
3	Problem formulation - Description and significance of the problem/phenomenon studied; review of relevant theory and empirical work; problem statement	131
4	Purpose or research question - Purpose of the study and specific objectives or questions	133
	Methods	
5	Qualitative approach and research paradigm - Qualitative approach (e.g., ethnography, grounded theory, case study, phenomenology, narrative research) and guiding theory if appropriate; identifying the research paradigm (e.g., postpositivist, constructivist/ interpretivist) is also recommended; rationale**	133
6	Researcher characteristics and reflexivity - Researchers' characteristics that may influence the research, including personal attributes, qualifications/experience, relationship with participants, assumptions, and/or presuppositions; potential or actual interaction between researchers' characteristics and the research questions, approach, methods, results, and/or transferability	135
7	Context - Setting/site and salient contextual factors; rationale**	133
8	Sampling strategy - How and why research participants, documents, or events were selected; criteria for deciding when no further sampling was necessary (e.g., sampling saturation); rationale**	133
9	Ethical issues pertaining to human subjects - Documentation of approval by an appropriate ethics review board and participant consent, or explanation for lack thereof; other confidentiality and data security issues	133
10	Data collection methods - Types of data collected; details of data collection procedures including (as appropriate) start and stop dates of data collection and analysis, iterative process, triangulation of sources/methods, and modification of procedures in response to evolving study findings; rationale**	135
11	Data collection instruments and technologies - Description of instruments (e.g., interview guides, questionnaires) and devices (e.g., audio recorders) used for data collection; if/how the instrument(s) changed over the course of the study	135

#	Reporting item	Page number
12	Units of study - Number and relevant characteristics of participants, documents, or events included in the study; level of participation (could be reported in results)	136
13	Data processing - Methods for processing data prior to and during analysis, including transcription, data entry, data management and security, verification of data integrity, data coding, and anonymization/de-identification of excerpts	135
14	Data analysis - Process by which inferences, themes, etc., were identified and developed, including the researchers involved in data analysis; usually references a specific paradigm or approach; rationale**	135
15	Techniques to enhance trustworthiness - Techniques to enhance trustworthiness and credibility of data analysis (e.g., member checking, audit trail, triangulation); rationale**	135
	Results/findings	
16	Synthesis and interpretation - Main findings (e.g., interpretations, inferences, and themes); might include development of a theory or model, or integration with prior research or theory	136-158
17	Links to empirical data - Evidence (e.g., quotes, field notes, text excerpts, photographs) to substantiate analytic findings	136-158; 180-189
	Discussion	
18	Integration with prior work, implications, transferability, and contribution(s) to the field - Short summary of main findings; explanation of how findings and conclusions connect to, support, elaborate on, or challenge conclusions of earlier scholarship; discussion of scope of application/generalizability; identification of unique contribution(s) to scholarship in a discipline or field	158-167
19	Limitations - Trustworthiness and limitations of findings	167
	Other	
20	Conflicts of interest - Potential sources of influence or perceived influence on study conduct and conclusions; how these were managed	N/A
21	Funding - Sources of funding and other support; role of funders in data collection, interpretation, and reporting	N/A

*The authors created the SRQR by searching the literature to identify guidelines, reporting standards, and critical appraisal criteria for qualitative research; reviewing the reference lists of retrieved sources; and contacting experts to gain feedback. The SRQR aims to improve the transparency of all aspects of qualitative research by providing clear standards for reporting qualitative research.

**The rationale should briefly discuss the justification for choosing that theory, approach, method, or technique rather than other options available, the assumptions and

limitations implicit in those choices, and how those choices influence study conclusions and transferability. As appropriate, the rationale for several items might be discussed together.

4.7. Chapter 4 - Appendix 2

Results Table with All Identified Themes, Theme Descriptions, and Key Quotations

#	Theme	Description	Key Quotation
Thematic Group 1: Prevailing approaches to data collection in HNA			
1.1	Household surveys are the default assessment method for most organizations	Surveys, particularly at the household level, are widely used as "quick shot" methods to provide a representative needs overview. There is a widely shared perception that RNA and MSNA need to be quantitative in order to be used.	"Yeah, I would say that, almost always, we're gonna have household surveys. And most of the time, there's focus groups. And then key informant interviews sort of vary depending on program design, and might be supplemented with like desk reviews, or more like open source sort of, you know, information that might inform things." (KI #03)
1.2	Key informant interviews are predominantly quantitative	Most KII are conducted with mostly or completely structured questionnaires in order to facilitate rapid data collection and analysis.	"So even KIIs, by definition, very short. … Those are, you know, structured KIIs, what you're calling KII at the household level. So a full-on qualitative, or purely qualitative approaches? I've not seen that. And even when they were they done-- doing them is one thing, using that data is a whole other story." (KI #17)
1.3	The role of quantitative data has increased due to KoboToolbox's widespread usage	The most commonly used tool is KoboToolbox. It allows for rapid data collection and analysis and has increased the relative availability and importance of quantitative data in HNA.	"But, you know, Kobo's been set up in such a way where it's so easy to get large volumes of data and into a … situation where you can process it and, you know, as yeah-- That day, you can do a survey and in the evening, get back--or even in the car on the way back from the site, you can start to look at the data almost immediately. It's, it's amazing." (KI #08)
1.4	OEQ are considered crucial for understanding complex subjects, but are mostly used with pre-coded options	Open-ended questions are considered useful for capturing nuanced information (such as capturing impact or the preferences of the population), particularly items that were not considered when designing the questionnaire. In surveys this is done by including pre-coded options in interview guides, often together with a free text response option.	"The open ended is where you get the true, like, impact. People might speak more to their unique situation. And I think that if we had some way to sort of analyze it at that level, we would just be able to speak more to our impact." (KI #03)

1.5	There is confusion about what constitutes qualitative and quantitative information	Quantitative methods, such as pre-coded open-ended questions in structured questionnaires, are used to measure qualitative information. The distinction between the two terms can therefore lead to confusion.	"You know, like, it's not a, it's not a quantitative question for me to ask you what your top priorities are right now, that's qualitative for you and your household. But because we collect it from, you know, you know, several thousand people and we present it in a graph, we make it into a quantitative thing. But it's still a qualitative piece of information." (KI #02)
1.6	Uncoded OEQ can be crucial for filling in information gaps	Uncoded OEQ in semi-structured or unstructured KII are seen as ideal to maximize the time of the respondent and fill in crucial information gaps.	"I think key informant interviews, especially when done towards the beginning really helps you also decide what it is. You know, it gives you It gives you a good background on what's going on where the needs are. And then it lets you cross check them with a number of people in a way that they… You know, you can ask for the specific questions, but then you also hear things that come out of those questions that may not be what you're asking and may further orient you towards the, you know, a deeper subject or something that you need to be looking at, at which you don't get that orientation in the same way, when you're just asking, you know, an A, B, C, and D question." (KI #12)
1.7	Paper and Excel are often preferred over CAPI tools for taking long notes in KII	Notes for what was said during semi-structured KII are often written on paper. These are often entered into preformatted Excel tables (sometimes called "transcript templates"). This is considered faster and more convenient than typing notes into a CAPI tool.	"They were doing fewer key informant interviews, and so I think it's just like easier for them to write down on paper. It was just easier. And then we didn't have to go through setting up a[electronic] form, find your form and then fill it out and then send the form." (KI #07)
1.8	Donors are exerting pressure in favor of quantitative methods, particularly MSNA	There has been a strong push from certain donor governments to increase the amount of quantitative data from household surveys, particularly through MSNA in order to be able to compare the situation in different crises and improve data-driven prioritization of funding. This has been driven in particular by the 2016 Grand Bargain commitments and led to the creation of the JIAF.	"I'm not sure if you're familiar with something called the joint intersector or analysis groups, … So there's been a big push on collecting household level information. And kind of like "MSNAs can solve everything", which is a bit of a . . . which isn't the case. Because you still need like, national GDP numbers that you're not going to get from a sample you have to kind of guess that's where you have to go everywhere." (KI #08)
1.9	Some organizations are biased in favor of	Some organizations / HNA professionals strongly prefer	"Management sees qualitative data as something that's easy to be cluttered if you want to have an easy approach. But I think the teams always rely only on

	quantitative methods	quantitative data, particularly household surveys, as the main method for understanding population needs.	household assessment data, because that is more scientific. That's, of course, the term that's often being used." (KI #18)
1.10	Qualitative methods are critical for answering "why" and "how" questions	Qualitative methods are widely considered to be critical for answering deeper questions that cannot be measured easily with quantitative methods.	"I mean, I think it's really useful in terms of understanding perceptions. … It gives you a much richer understanding, it really enhances any understanding that you might get from a purely quantitative viewpoint." (KI #12)
1.11	Qualitative and quantitative methods are seen as complimentary	Household surveys, FGD, and KII were cited as complementary methods for answering different questions during HNA.	"It's just that when we don't use qualitative data, we might really miss out on really key elements. And also, if we miss out on having the quantitative component, we also miss out on representativity, which is like also super important, right? So in an ideal case scenario, we have both Yeah. Both aspects can complement each other." (KI #13)
1.12	Qualitative skills are slowly increasing	There is a growing demand for qualitative research skills in the humanitarian sector, including anthropologists.	"I feel like if there's a trend that I'm seeing in humanitarian work at the moment, like especially over the last, you know, 12 months to two years, is somehow a greater, a greater desire to be more context specific." (KI #02)
Thematic Group 2: Constraints related to quantitative interview methods			
2.1	Some information by respondents is never recorded	Some potentially important information given by respondents is lost because of structured data entry methods.	"If someone asks a question … 'Do you have access to the markets?' 'No … I got intercepted by military, I tried to do this, and the roads were like' … That's interesting. And it's not there." (KI #07)
2.2	For some sectors, household surveys are not the best method to obtain important information	Household surveys are not considered useful for some humanitarian sectors. KIs cited in particular health and protection as two sectors that require qualitative methods.	"[I am] trying to stop people from asking diagnostic questions in household surveys because people are not great at self-diagnosing. It's a general rule. And even that data is really unreliable for us. We can't use it for anything, it's not operational. So, there are some things that we collect from MSNA type surveys, but up until very recently, and I would say maybe this upcoming year, the bulk of it has been one question on 'How far away is the nearest health facility in minutes?'" (KI #05)
2.3	Some groups are suspicious of quantitative assessments, particularly remote survey methods	Certain populations (especially rural, but also specific groups such as displaced Rohingya) have been found to react suspiciously to quantitative methods, leading potential to higher rates of refusal or not being as forthcoming with	"They've reported it 100 times, they don't like quantitative data collection, especially when they're asking about people's behavior and how they're feeling and things like that. They don't really quite understand why someone would want that information. They don't find that consultative, they don't find that any type of conversation. So it's hard to get any type of meaningful data if that's not what they would like to communicate through." (KI #10)

		their responses. This is true in particular with remote survey methods (CAPI).	
2.4	Structured KII are overused	Many HO conduct KII using structured interview guides even more qualitative approaches are possible, either out of habit or for convenience reasons.	"So we're collecting data that way because it's the only way we can collect data at the speed we need to collect it. In the immediate aftermath of an event, so we know our data is questionable. We were sacrificing immense quality, just to get it fast and make sure that we've got something with which to base the next round on, you are still supposed to be collecting better data. A few weeks down the road like this should not be the basis for everything. It's just the initial planning. So we know it's questionable then the problem is when you're still using those methods when you no longer need to. And that happens far too often just because they're easy. So effectively quick and dirty has its time in place. We need to be careful of not overusing it and we do we overuse it a lot" (KI #05)
2.5	A lot of quantitative data is contradictory or is not explained	In some emergencies there can be a lot of quantitative data from multiple assessment rounds—at times with conflicting results. Despite this wealth of data, there is sometimes little actual understanding of the population, which can only be achieved with good qualitative data.	"You're covered in data. But it took me I think, three or four months before I was like, wait a minute, I'm drowning in all this secondary data. But I don't feel like I understand the population at all. And all I feel like I'm seeing is conflicting quantitative data that says this, many people need this, but then they're doing this behavior. And I just was, I think I spent a good two weeks running around talking people trying to understand because I just assumed that I was missing something." (KI #10)
2.6	Limited access to affected population can bias assessment results	Access to KI or households from the affected population can be very constrained in emergencies, either due to logistics or limitations imposed by political or security actors, introducing potential biases to the HNA results.	"The most so they were able to do basically a lot of it was through key informant interviews, both through the humanitarian community, and I would say including national Local authorities, and then they were able to speak with some representatives and I would say representatives because not all of them were official local, like local officials within some of the key targeted geographic regions… And where we were able to have access, the government is quite controlling, over access within the country. … But we weren't able to speak to the population, we weren't able to do any kind of household interviews or even, you know, like, have any kind of group discussion of any kind, not even sort of informally, with either displaced or host communities." (KI #12)
2.7	Household surveys can put respondents and staff in danger	Household surveys can put respondents and staff in danger in conflict areas or cities with a high incidence of targeted violence.	"In Honduras, for example … it was way too sensitive and dangerous to run a survey, it was too dangerous to enter the communities with a big group of enumerators running around with their, you know, tablets, and that's just super dangerous. So, what happened there, we mainly worked only through focus group discussions with local NGOs or partners who had been working within the communities for years before" (KI #13)

2.8	Representative household surveys are not appropriate during Phases 1 and 2	Household surveys are not appropriate right after a sudden-onset emergency due to access and sampling challenges.	"The term rapid needs assessment has a lot of different views. Some people view that as a household level representative assessment, that needs to be done in two weeks. Which is wrong, it can't happen. " (KI #14)
2.9	Questionnaires are not contextualized enough	Although often being translated, survey instruments are often not contextualized with local team. This step ensures that no relevant questions are left out or that questions are culturally sensitive.	"Many people think contextualizing tools is translating it. What is not [done is] sharing with a team and trying to formulate the questions to see if they are culturally sensitive or not, what needs to be changed, what additional questions need to be made." (KI #20)
Thematic Group 3: Constraints related to qualitative interview methods			
Cluster 3.1: Capacity limitations			
3.1.1	Qualitative analysis is too time consuming for emergency contexts	Proper analysis of qualitative data requires a lot of time. That makes it unsuitable for most emergency contexts where results need to be established quickly.	"In sudden onset disaster, my experience and my... I don't have a vision for the future of qualitative data in sudden onset disaster. It's too complex. They struggle too much to do what already they should be doing, which is SDR, KIIs and direct observation. If you add in there anything which is text... We do not recommend it, even now. In training, I'm used to saying, "Hey, do not collect focus group data in a sudden onset disaster. You can do this in phase three. You can do this after one month. No problem. In the first month, guys, you need to be quick, so it's more important to be fast than to know the little nuances and differences between the gender and the old and the young and so on." (KI #23)
3.1.2	There are not enough staff with expertise in qualitative methods	Organizations often lack sufficiently skilled staff to analyze qualitative data. In many organizations, staff asked to support or lead HNA have trained almost exclusively in quantitative skills.	"I think it's often easier with an emergency team, because you have people that it's a smaller group people, you train them. And usually they're the people that are coming with quite a lot of experience. And that doesn't mean that they always do things, know how to do things well or not, but like they're coming with a certain level of experience. Whereas at country program level or even with enumerators you might hire on the spot. Often it's not necessarily those people with the same level of experience. " (KI #12)
3.1.3	Interviewer training is much too short	Interviewers involved in data collection for HNA often do not have sufficient skills for qualitative methods.	"So in some cases, if they really have done a day and covered a lot of what they should, then I'm not necessarily thrilled, but I am happy because at least had done that because there have been like, would there have been cases in the past where we would do like a one-hour training, and that that is the reality of NGOs sometimes. Quality has come a long way in the past few years. And one of those areas is needs assessments. And so a day is already an improvement on what was the case in some areas, not saying it's right, I'm just speaking the truth." (KI #14)
Cluster 3.2: Notetaking / transcript challenges			

3.2.1	Manually created notes are often of poor quality	Manually written notes of what a respondent said are often of very poor quality. Contributing factors are insufficient training of the interviewer as well as the difficulties of entering long notes into CAPI tools.	"Every time we tried it was a total disaster [laughs] it was a total failure. It requires so much. … Are they transcribing properly? Are they interpreting what the person is saying? They are doing their eighth interview today. They're just talking whatever. They're tired. So, it's never actually worked really. We tried that way." (KI #21)
3.2.2	Verbatim transcripts of interview responses are only rarely created because audio is almost never recorded	Verbatim transcripts of interview responses (created later on based on an audio recording) are very rare in HNA. The main reason given is the amount of time needed to create the transcript, creating a translation, and then condensing or coding it.	"Audio? I don't see that. I have yet to see that in an operation that I've worked in." (KI #16)
3.2.3	Audio files can create data protection risks or make respondents more reticent	Audio files created during interviews were seen as adding a potential layer of risk to the respondent due to the lack of anonymity.	"We often try not to record much just because from experience, people are very nervous about recordings. And so especially if we're talking to like people in the affected population, so it is an option to record. But we guide our teams to only do that if they really know people are comfortable." (KI #14)
Cluster 3.3: Language issues			
3.3.1	Significant language gaps hamper qualitative data collection and analysis	There is a gap between the languages spoken by affected communities and those of of the responding humanitarian organizations. This significantly hampers the ability to collect and analyze qualitative data.	"We had to use also translators and interpreters because sometimes we had people that spoke only English and then we find one translator from English to Portuguese and another one from Portuguese to the local language. Ideally, we were trying to get someone that could go from English to the local language, but sometimes it was not possible." (KI #20)
3.3.2	Notes are rarely translated properly	Notes taken from open-ended interview responses are often not translated into a language spoken by the person analyzing the data due to a lack of time or resources. As a result, they are often not considered during the analysis.	"The problem is, as I said, I mean, even just "others", most of the time, [the notes] don't get translated. … And by then now, because they're so used to the systems where you can collect that on KoboToolbox, and then it goes straight into whatever automated system you've created, and you've got a graph, and you have your results already, they don't want to wait. Those steps slow the process down. And they also are extremely costly." (KI #05)
3.3.3	Meaning is lost because analysis is not	Local staff who can contextualize responses or clarify	"Why don't we show the qualitative data so that it must be in the native language? … We must be able to do it in … Swahili, in Farsi, in Somali—not

	done by or with local staff	misunderstandings are often not involved.	just in English, French, Arabic or Spanish. That's not good enough. There are so many words that change the entire meaning, or it will not exist in a certain vocabulary, that can completely screw your assessment. " (KI #21)
Cluster 3.4: Systemic challenges			
3.4.1	Qualitative and quantitative methods are often not integrated	Qualitative data collection is often done by completely separate teams in the same organization; or it is done instead of quantitative methods in order to maximize resources.	"Frequently, what happens—because assessments are multifactorial—is we wind up shoving more of certain sectors into the qualitative, and some of the others out into the quants data. So, we can ask most economic type questions in quant data. And so most of that gets sidelined and pushed out of the qualitative data, which creates problems sometimes." (KI #04)
3.4.2	There are systematic biases against qualitative data	Biases against qualitative methods can be found at all levels, ranging from assessment specialists and senior leadership up to the organizational level.	"Qualitative interviews … have kind of pointed to deficiencies in the quality of food assistance, for example, serious deficiencies have been brushed aside by key organizations involved in food, because it well, "Just because 12 households or whatever told you that. That doesn't really mean anything in a response of one million people. Our teams who work for us, the organization providing the food, found that there wasn't a problem. So it's fine."" (KI #02)
3.4.3	The trend to collect data remotely reduces the ability to collect qualitative data	There has been a drive to conduct assessments increasingly through remote data collection methods, a trend accelerated during the COVID-19 pandemic. This has reduced the amount of qualitative data collected.	"There are risk associated to do remote operation. For sure. And the more we do, Yes, there will also be strategy of adaptation from the from the beneficiaries, strategy adaptation from the communities that, you know, It's like they understand how to reply, and I say, yeah, you know. You know better than me that you can go in certain communities in Congo, and basically, the teacher says Okay, what do you need? What do you need? A list of women-headed household or with disability? The other one, for IDPs: for less than three months of the IDPS more than three months? Everything is ready, right? Because they already got exactly how to get from the system. And I think it's fair. It's totally fair. And but that's going to happen. Also, we do remote data collection more and more and more of that repeated thinking. And per se, again, is not a total tragedy. If we are honest enough with what knowledge does bias to, to factor this in our analysis, to maybe identify possible risk mitigation measures that we can use, etc. " (KI #21)
3.4.4	Donors' short funding cycles don't allow for qualitative methods	Humanitarian organizations may be hard pressed to conduct a needs assessment within a very short period due to a donor's short application deadline. This limits the use of qualitative methods.	"The time sensitivity that comes from the donor that says, Yeah, we are opening a call and you have 10 days to apply. And you want to do a need assessment, because you want partly to feed the donor's demand for fresh data, partly because you actually care. And you think it would make a better program if you were to do a little need assessment. But if the donor wants it in 10 days, and you have to do the analysis, and you have to do the proposal writing that there is it is urgent, even though it didn't have to be" (KI #19)

3.4.5	The push for more secondary data analysis and automation cannot replace qualitative primary data	Challenges with using automated coding of qualitative data include the errors made by existing tools, lack of trust in the results, potential lack of human verification.	"And now there's a push to automate analysis—I really loved when I heard that one: DEEP is going to automate analysis, apparently, yeah, yeah. … But so the that push, I think that we need to check that we're right. If you're going to do those sorts of things, if you're going to rely on algorithms, you need to check that they're doing what you think they're doing, that they're telling you what you think they're telling you. And the only way I know how to do that is with qualitative information." (KI #05)
Thematic Group 4: Challenges specific to pre-coded open-ended questions			
4.1	Pre-coded responses to open-ended questions in questionnaires can lead to false data	Pre-coded responses for recording responses to open-ended questions in household surveys are problematic because the response options are often not contextualized for the local context. This risks that relevant information is not captured and programming is informed by inaccurate results.	"The ideal needs assessment [at the UN level] is a household survey. And it's quantitative. But they're collecting qualitative data in a quantitative format. And we make these ridiculous assumptions that we know the answer options we should be giving without testing them, and we roll them out globally and keep using the same ones. And because we use them in so many places, we consider that evidence that they work. But we never often go in and actually check that that's what we should be asking. So there's problems with how we phrase them. There's bias built into the questions we use, that we repeatedly implement. " (KI #05)
4.2	OEQ are considered an obstacle to obtaining large samples	OEQ, whether with or without pre-coded responses, are often considered as an obstacle to collecting data from larger samples. This is due to the additional time needed for respondents to answer such questions, as well as the increased time required to analyze responses that were entered as text.	"We try to have very few open ended because we want, we're trying to get larger numbers of responses. And to get a better picture of more people. And, it's just really hard. Often in a needs assessment, you're working before you have like a full program team. So, it makes it harder to do analysis, the heavier the tool it is." (KI #03)
4.3	Most text responses to OEQ are not properly analyzed	Qualitative text responses to OEQ are not systematically analyzed, or not analyzed at all, for a large proportion of HNA. KI underline the importance of doing this properly but most immediately qualified this to say that manual analysis (and often translation) of these notes is either not done well or not done at all.	"People do type in the text. But if I'm very honest, I don't think it's being used as much as it could be, because you focus on your 90% of the data that's easily processed. Then you'll say "I'll come back to that," and then you don't come back to that." (KI #09)

4.4	Interviewers may not choose the correct responses among pre-coded options	Interviewer may not select the correct responses among pre-coded options, either for a lack of familiarity with the terminology, personal biases, or fatigue.	"That's why sometimes we even don't use the "other". It really depends, I guess. So. It can be a tricky thing. That also really [depends on] how well you trained. And you're right. So… Because we don't want to end up with a survey where they always use "99 - Other". Yeah. Wow. Okay. Great. Thank you for that." (KI #13)
Thematic Group 5: Suggestions for improving the role of qualitative interview methods in HNA			
5.1	More funding is needed for longer training sessions	Interviewers require more training to be able to do more qualitative data collection. Funding was universally identified as the limiting factor to conducting longer training sessions.	"You need much more rigorous training. Right now trying to get three days of training for enumeration is a stretch for qualitative data, you would need way more than that. And it's just that timeframe. And funding is a major problem." (KI #05)
5.2	Ensure assessments teams have qualitative expertise and speak local languages	Assessment teams should include people who are trained in qualitative methods, if need be, by hiring additional staff or training existing ones. Involving local staff in the design and analysis stages is essential to ensure as little meaning as possible is lost to translation.	"We have a massive problem with assessments because we're not hiring the right people to do them. So we're focusing on that IMO type, who don't have the right background… Even if you want people with serious humanitarian background, hire your M&E people to run the assessments. … Make sure that you've got some people who know how to design it and train people on qualitative data collection, because it's a lot harder to do." (KI #05)
5.3	MSNA should contain more qualitative methods	In protracted emergencies, MSNA should be reduced or replaced with qualitative methods.	"For protracted crises, I do not understand why we are putting so many questions in our questionnaires. … It's insane. 450 variables. 900 variables. It's crazy. It's hours of discussion. It's bullshit. Everything that you ask after 45 minutes, it's bullshit. The guy just wants to leave or to be somewhere else. … I think it's not smart. It's simply not smart what we are doing right now. Putting 1.5 million into a field assessment, there is barely a secondary data review done beforehand. There are barely any qualitative interviews being done in advance to help shaping the questionnaire—which goes for everything in this stupid questionnaire. For me it's a real waste of money and I find it damaging, harmful, non-ethical, irresponsible. … In the field, we should have much more qualitative data." (KI #23)
5.4	Conduct hybrid surveys with a mix of quantitative and qualitative methods	Combining quantitative and qualitative approaches in surveys could generate data that is timely, representative, and provide more depth.	"You know, what, if I can have a shorter questionnaire, and then I ask my two typical things, right? And then I tell them, okay, which are your three priorities? Right, the classic question. "Which are your three priorities?" And the person says "food." And then you tell them "Okay, great. Can you explain me why?" I could probably cut my survey in half. Okay. "Which are your priorities right now?", and then This, this and this, and this. "Okay, why? Tell me." And then I

			can analyze the why. I could get rid of at least 20-30 questions, easily." (KI #21)
5.5	More audio recordings would be useful to capture all details	Having audio recordings of interview responses would be beneficial for multiple purposes. This includes the ability to do quality control of interviews, derive more data from a response, keep the interview shorter, and to create transcripts.	"I would, if I could choose, I would always go for transcripts and recordings." (KI #01)
5.6	Automatic transcription would allow more widespread qualitative data collection	Automatic speech recognition (ASR) of interview responses would allow organizations to conduct more qualitative interviews. ASR would reduce the amount of time spent by local teams. But gaps exist regarding the languages required.	"If we had a way to take qualitative information, and analyze it to the depth of what we can do with quantitative information, like, and we could interview I mean, and we could interview like 10 times as many people as we currently do. And we could then have far more representative data about a population at a much more intimate and deeper level." (KI #03)
5.7	Machine translation would be helpful for systematically translating transcripts or notes	Translating interview transcripts or manually entered notes would help generate useful insights for HNA. Machine translation (as well as streamlined reviews by professional translators) would be needed to obtain results quickly enough.	"I think if you had like, Google Translate, and you could export the qualitative wording … We would copy and paste that into our Excel file and start doing our analysis that way around. … But I think the biggest hurdle is the translation." (KI #14)
5.8	Automatic coding would be helpful to speed up analysis	Automated methods for coding qualitative data would save a lot of time and allow for using such data more widely.	"I think if there was a way that the kind of initial coding or sorting of that data could be done in an automated, quick way, so that it relieved the fact that it just takes so long to do it manually. And I think that would be kind of the big, the big lightbulb, the big key that would change everything." (KI #02)

4.8. Chapter 4 - Appendix 3

Interview Guide

1. To start off, could you please describe your current role at [ORG], what types of responsibilities you have, and what kinds of data collection are you typically involved in?
2. I'd like to first understand the role of evidence that's used to decide what kinds of assistance [ORG] will provide where, when, and to whom. How would you describe the role that evidence plays to inform [ORG]'s decisions? What kinds of decisions do your needs assessments typically need to inform?
3. What would you say are the greatest challenges related to understanding affected people's needs?
4. I'd like to understand how you and [ORG] in general go about understanding people's needs. Could you talk me through how do you usually conduct needs assessments, both before and since the beginning of the pandemic?
5. Do you think that some of these changes since the beginning of the pandemic are here to stay?
6. I'd like to talk specifically about primary data now. What role do qualitative methods play in your needs assessments when collecting primary data from affected people?
7. Do you consider open-ended questions a type of qualitative data?
8. How do you collect qualitative / open-ended data?
9. How do you typically analyze qualitative data (transcription, translation, keywords)?

10. Have you ever collected audio as part of a form?

    IF YES: How do you and your teams usually deal with audio responses?

11. I'd like to switch over to ethics now. Are there any ethical issues that keep you up at night related to your work? Are there major concerns?

12. Generally, how do you ensure that the data you collect is processed in a way that doesn't pose risk to your respondents?

    [PROBE: What about informed consent, anonymization, etc.]

13. Are there any particular data protection guidelines or similar framework that you use or that you need to adhere to in your work?

14. What role do you think AI or algorithms will play in humanitarian response in the future? Do you think they pose any additional risks for use in humanitarian response?

15. Do you have any other comments you'd like to share? Do you have any questions for me?

Chapter 5 Ethical Implications Related to Processing of Personal Data in Humanitarian Crises: A Scoping Review

Abstract

Humanitarian organizations are rapidly expanding their use of data in the pursuit of operational gains in effectiveness and efficiency. Ethical risks from the processing of such data, including through the use of artificial intelligence, are increasingly recognized but are insufficiently addressed by existing humanitarian data protection guidelines. This study reports on a scoping review that maps the range of ethical issues that have been raised in the academic literature regarding data processing of people affected by humanitarian crises. We systematically searched databases to identify peer-reviewed studies published since 2010. Standardized forms were used to extract data and compare findings and study characteristics. The ethical issues identified in each study were grouped into the ethical value categories of autonomy, beneficence, non-maleficence, and justice. The study protocol followed PRISMA-ScR reporting guidelines for scoping reviews, and yielded 100 relevant studies after screening 8,387 papers. The majority were published after 2015 (n = 73) and included at least one author based in a high-income country (n = 93). Most studies focused on processing data for assessments (n = 36); discussed the use of social media (n = 48), and identified ethical issues based on real-life examples (n = 54). Only eight included an author from a lower-middle income country while none included an author from a low-income country. One in four (n = 25) discussed technologies related to artificial intelligence. A total of 22 ethical issues were identified. Published research shows widespread concerns that data processing in humanitarian assistance can cause additional harm, may not provide direct benefits, may limit the autonomy of affected populations, and has the

potential to lead to the unfair distribution of resources. Specific and robust guidance is needed to navigate emerging ethical concerns. Further empirical research is needed into ethical issues arising from the use of commonplace humanitarian tools and data management processes. In addition, case studies of early adoptions of AI in humanitarian operations should address which ethical considerations are being applied.

Ethical Implications Related to Processing of Personal Data in Humanitarian Crises: A Scoping Review

Organizations involved in humanitarian assistance currently face significant pressure on multiple fronts. Protracted violent conflicts in many countries, intensified by worsening conditions due to climate change, have led to 82 million forcibly displaced people, and more than 235 million people need some form of humanitarian assistance (OCHA, 2021; UNHCR, 2021). In addition, the gap between humanitarian needs and available funding reached its highest level on record in 2020 (OCHA, 2021). The aim of this review is to map the range of ethical issues that have been raised in the academic literature regarding data processing of people affected by humanitarian crises.

A review by Schofield et al. (2021) found that the vast majority of included studies discussing "ethical challenges" in healthcare had failed to include an explicit definition of how that term was understood by the respective authors, leading to potential misunderstandings and ambiguity. This section, therefore, will first provide working definitions for the key terms and concepts discussed in this study. Humanitarian assistance is understood here to refer to coordinated actions that save lives and alleviate suffering of crisis-affected populations (UN OCHA, 2003). It also includes "protection", which "encompasses all activities aimed at obtaining full respect for the rights of the individual in accordance with the letter and the spirit of the relevant bodies of law" (IASC, 1999, p. 4). Humanitarian crises are defined here as a "series of events representing a critical threat to the health, safety, security or wellbeing of a community, usually over a wide area" (WHO, 2007, p. 7). For the purposes of this study, data processing is understood as: "Any operation or set of operations which is performed on data or on sets of data, whether or not by automated means, such as collecting, registering, storing, adapting or altering,

cleaning, filing, retrieving, using, disseminating, transferring and retaining or destroying" (OCHA, 2019, p. 47). Ethical issues are defined in this study as actions that may not conform to moral standards, particularly those set out by various humanitarian principles (Slim, 2015, pp. 47–145) because of the risks they present.

Humanitarian organizations rely on processing increasingly large amounts of data to inform their operations, much of which is collected directly from affected populations (e.g., through registrations, household surveys, or cash disbursements). At the same time, the people working for these organizations have themselves often become targets of kidnappings and killings, which has led organizations to increasingly resort to remote methods of managing operations and collecting data from affected people (Donini & Maxwell, 2013; Humanitarian Outcomes, 2021). The COVID-19 pandemic has accelerated this trend of the increased use of remote methods. This combination of factors has led to an exponential increase in the amount of personal data that is being distributed, stored, and analyzed in various locations around the world. At the same time, humanitarian organizations are continuously seeking innovations involving information and communication technologies (ICT) in the pursuit of operational gains in effectiveness and efficiency.

Organizations turning to new or existing digital tools to collect, store, or analyze data more efficiently may knowingly or inadvertently introduce new ethical issues affecting people who are already vulnerable (International Committee of the Red Cross and Privacy International, 2018). Weighing the responsible use of new technologies in humanitarian crises is fraught with a number of ethical issues (Sandvik et al., 2014) that are increasingly being highlighted in specific circumstances such as refugee registrations (Jacobsen, 2015), health emergency response (Perakslis, 2018), or the use of drones in humanitarian assistance (van Wynsberghe & Comes,

2020). In practice, ethical decisions are made—knowingly or unknowingly—on a daily basis about what data to collect, which tools to use, or how and with whom to share this information to avoid adverse consequences (Raymond et al., 2016; UN Global Pulse, 2016). In light of such challenges, organizations rarely choose to forego new tools altogether, such as Oxfam's decision in 2015 to halt the use of biometrics in its programs in order to assess the potential risks (The Engine Room and Oxfam, 2018). Rather, some organizations are more likely to invest in new innovations without considering, weighing, or fully grasping the long-term ethical issues (Parker, 2019).

However, because of these challenges, more guidelines are now being produced for the ethical processing of data for humanitarian assistance purposes, with the goal of minimizing or eliminating risks to vulnerable people. Notable examples include *Data Responsibility in Humanitarian Action* by the Inter-Agency Standing Committee (2021), the *Handbook on Data Protection in Humanitarian Action* by the International Committee of the Red Cross (2020), and the *Data Responsibility Guidelines* by the United Nations Office for the Coordination of Humanitarian Affairs (2019). Similarly, regulatory environments are changing in many countries (such as the European Union's General Data Protection Regulation, GDPR), which have moved many humanitarian organizations to change their approaches to data processing in order to improve data privacy. Focusing on the issue of ethical design of new tools, Krishnaraj et al. (2021) have created practical guidelines that aim to mitigate risks as early as possible (Elhra & Humanitarian Health Ethics, 2021). But the speed of technological innovation means that such guidance can quickly become out of date as new data technology tools appear and organizations respond to new circumstances (such as insecure environments or lack of access to populations during the COVID-19 pandemic). Artificial intelligence (AI) systems that use machine learning

and other methods for automating data processing may usher in a completely new set of ethical issues that humanitarian organizations will have to confront (Tegmark, 2017). Although a considerable number of studies discuss the ethics of using various technologies in humanitarian assistance, to date, there is little evidence that there has been a comprehensive review of relevant ethical issues in the published literature.

Conducting this type of review is challenging due to the wide-ranging nature of humanitarian assistance, lack of well-defined nomenclature for data processing technologies and activities, and that relevant research may be published in the intersecting fields of ethics research, design, engineering, health, medicine, geography, development, social science, and technology research, among others. Previous scoping reviews focusing on humanitarian assistance only addressed more limited contexts, such as natural disasters (Tansey et al., 2018), and displaced populations (Makhoul et al., 2018), or did not include terms to capture more novel humanitarian activities such as responding to large-scale migration or public health emergencies (Pal et al., 2019). Other relevant studies in the past have included a literature review focused on social media and privacy issues (based on literature published between 2013 and 2014 (Watson & Rodrigues, 2018), a scoping review on the types of digital tools used during the 2014-2016 Ebola outbreak in West Africa (Bempong et al., 2019), a scoping review on the impact of health-related tools in humanitarian crises (Mesmar et al., 2016), and a scoping review on ethical considerations related to the use of drones in humanitarian assistance (Wang et al., 2021). However, we have not found a sufficiently comprehensive sets of keywords that could be used to search databases for any of this study's three inclusion criteria (people affected by humanitarian crises, processing data for humanitarian assistance, meaningful discussion of ethical issues).

Another distinct challenge is the lack of established ethical categories or theories used by studies discussing ethics in the humanitarian sector (Wang et al., 2021). First introduced by the International Federation of Red Cross and Red Crescent Societies (Pictet, 1979), the four humanitarian principles (humanity, impartiality, independence, and neutrality) are now widely used among many humanitarian organizations (see, for example, MSF, n.d.), in international law (OCHA, 2012), as well as in ethical codes attempting to guide the actions of the humanitarian sector as a whole (see, e.g., IFRC, 1994; Sphere Association, 2018). However, previous studies have shown the difficulty of applying these humanitarian principles in everyday practice (Hilhorst & Schmiemann, 2002), in guiding the use of information technology (Raymond & Card, 2015), or in mapping humanitarian organizations' ethical obligations (Broussard et al., 2019). In particular, the broad *humanity* "principle" has been argued as being better understood as an absolute moral value rather than an ethical principle (Slim, 2015, p. 62). As a result, a growing number of studies use the ethical value categories of autonomy, beneficence, non-maleficence, and justice, which are widely used in the fields of bioethics and research ethics (Beauchamp & Childress, 2019), as a better operational ethical terms to reference ethical issues inherent to humanitarian practice (Cawthorne & Wynsberghe, 2019; P. N. Pham & Vinck, 2012). This study therefore uses these four ethical value categories to group the ethical issues identified in the literature and to better link the nascent field of humanitarian ethics to the larger theoretical and practical advances in the fields of bioethics and research ethics. A definition of each ethical value category is provided in Table 11. For the purposes of this study, the ethical value category of beneficence is considered to include the humanitarian principles of independence and neutrality as these can be best understood as "mechanisms" (Orbinski, 1994, p. 7) of the

imperative to reach and provide benefits to the greatest number of affected populations in conflict settings.

The aim of this review is to map the range of ethical issues that have been raised in the academic literature regarding processing relevant to people affected by humanitarian crises, and to provide suggestions for future research.

Table 11

Definitions of Each Ethical Value Category, Based on Beauchamp and Childress (2019, p. 16)

Principle	Definition
Respect for autonomy	Respecting the decision-making capacities of autonomous persons
Beneficence	Providing benefits and balancing benefits against risks and costs
Non-maleficence	Avoiding the causation of harm
Justice	Distributing benefits, risks, and costs fairly

5.1. Methods

5.1.1. Study protocol

We chose to conduct a scoping review as this method is best suited for generating a broad overview of relevant evidence, to examining emerging areas of research, to clarifying key concepts, and to identifying gaps in the literature (Peters et al., 2015). A study protocol was developed prior to data collection and screening using the scoping review method established by Arksey and O'Malley (2005) as further refined by Levac et al. (2010) and follows the framework maintained by the Joanna Briggs Institute (Peters et al., 2017). The protocol was revised based on feedback received from the research team and incorporated the results from a pilot conducted for this study November-December, 2019. It follows the Preferred Reporting Items for

Systematic Reviews and Meta-analysis for Protocols—Extension for Scoping Reviews, or PRISMA-ScR (Tricco et al., 2018). The final version of the study protocol is available in Chapter 5 - Appendix 1.

5.1.2. Identifying the research question

The specific research questions of this scoping review were:

1. Which ethical issues have been raised in the literature related to processing data from people affected by humanitarian crises in order to inform humanitarian assistance?
2. To what extent do real-world examples of ethical issues reflect the concerns presented in the literature?
3. Which technologies were the focus of concern over these ethical issues?

5.1.3. Eligibility criteria

The following eligibility criteria for the selection of relevant studies were established *a priori* as per the categories and requirements for scoping review protocols (Tricco et al., 2018).

- *Condition/Domain*: Ethical issues stemming from the processing of data relating to people affected by a humanitarian crisis with the explicit goal or potential of informing humanitarian assistance.
- *Population*: People affected by a humanitarian crisis, including armed conflicts, natural disasters, and large public health emergencies, as well as refugees and transborder migrants fleeing from such a crisis—regardless of their current location. We also included studies that concern humanitarian assistance (including related fields such as disaster response or emergency management) that are global in scope. Studies about

natural disasters were only included if the study focused on events in low or lower middle-income countries (defined as countries that ranked low income or low middle income at least once by the World Bank between 2011-2020) (World Bank, 2020). The Ebola outbreak in West Africa (2014-2016) was included as it was widely considered to be a humanitarian crisis in scope (UN News Centre, 2015). We used the Financial Tracking Service by UN OCHA (n.d.) to judge if an event should be considered a humanitarian crisis (defined as whether a given country was a recipient of humanitarian aid in the same year).

- *Interventions*: Data processing relating to people affected by a humanitarian crisis with the explicit goal or potential of informing humanitarian assistance. Excluded were studies that focus on technologies that do not process data on affected people, such as robotics for clearing debris or land mines, algorithmic models for predicting the occurrence or impacts of natural hazards, or tools used for planning humanitarian logistics (e.g., relief/distribution networks, supply chain management, and resource scheduling).
- *Outcomes*: Studies that investigate ethical issues stemming from the processing of data (as defined above) were included only if they contained a significant discussion about this subject. During the screening stage, studies were eligible for inclusion if the abstract referenced or mentioned potential ethical issues. During the full text review this was assessed qualitatively by two reviewers.
- *Study Designs*: All study designs were eligible for inclusion, including empirical studies, commentaries, and theoretical papers. Excluded were non-peer reviewed studies as well as book reviews.

- *Context*: For feasibility reasons, we restricted the review to studies published between 1 January 2010 and 31 December 2019.
- *Setting*: Studies in all countries or territories affected by a humanitarian crisis (or relevant host countries for refugee or cross-border migrant or displaced populations) were included, as defined above.

5.1.4. Search Strategy and Information Sources

Comprehensive literature searches of electronic databases were conducted on 31 March 2020, using Ovid, Ebsco, Web of Science, and Proquest to search 20 databases for relevant studies. Only studies published in English, French, or Spanish were included.

As recommended by the scoping review guidelines described above, keywords were selected and piloted in multiple iterations to identify all relevant articles. The authors had previously identified 34 studies, and these were used as a minimum search target. After an initial search showed that only 13 were included, we repeated the database search over several iterations with additional terms until all 34 studies were reflected in the results. This yielded additional keywords such as “risks” and “challenges” to represent ethical challenges, as well as “innovation” and “experimentation” which are sometimes used to refer to data processing activities. Further, careful searching for terms such as “acute malnutrition” or “forcibly displaced population” were also found to describe specific phenomena in a humanitarian crisis without using terms such as “refugees” or “humanitarian” in the study’s metadata. Likewise, to find all studies that discuss processing data of affected people, we iteratively expanded our search terms to include specific technologies (e.g., biometrics, remote sensing), emerging practices (e.g., remote management, crowdsourcing), or shorthand keywords introduced by researchers (e.g., experimentation, crisis informatics, innovation). A sample of the search strategy for the OVID

databases is displayed in Table 12. The complete search syntax for each database can be found in Chapter 5 - Appendix 2.

Table 12

Search Strategy for OVID Databases

Concept		Keyword and syntax
Humanitarian assistance	1	humanitarian*.tw.
	2	relief work.tw.
	3	aid work.tw.
	4	(disaster? Adj (relief or response? Or assistance)).tw.
	5	emergency relief.tw.
	6	((conflict? Or war?) adj10 (human rights or public health)).tw.
	7	(ebola adj6 (west africa or sierra leone or liberia or guinea or 2014 or 2013)).tw.
	8	acute malnutrition.tw.
	9	(refugee* adj2 (camp* or assistance or population?)).tw.
	10	(displace* adj2 (forced or forcibly or population? Or human? Or internal*)).tw.
	11	(((population? Or person* or communit*) adj3 affected) adj1 (conflict? Or violence)).tw.
	12	or/ 1-11
	13	(cris?s or emergenc* or disaster?).tw.
	14	humanitarian*.af.
	15	13 and 14
	16	12 or 15
ICT for data collection	17	ict.tw.
	18	technolog*.tw.
	19	((data or information) adj2 (system* or manage* or collection or analys?s or process*)).tw.
	20	(blockchain or distributed ledger).tw.
	21	(ai or artificial intelligence or machine learning or algorithm*).tw.
	22	biometric*.tw.
	23	smartphone app*.tw.
	24	remote sensing.tw.
	25	analytics.tw.
	26	digital*.tw.
	27	experimentation.tw.
	28	automat*.tw.
	29	innovation?.tw.
	30	remote management.tw.
	31	cyber.tw.
	32	big data.tw.
	33	(sms or text messag* or interactive voice recognition or online survey*).tw.
	34	(kobotoolbox or kobo or odk or open data kit).tw.
	35	crowdsource*.tw.
	36	social media.tw.
	37	crisis adj (informatics or data or map*).tw.
	38	digiti?ation.tw.
	39	datafication.tw.
	40	or/ 17-39
Ethical concerns	41	concern?.tw.
	42	risk?.tw.
	43	challenge?.tw.
	44	harm?.tw.
	45	privacy.tw.
	46	protection?.tw.
	47	humanitarian adj (principle? Or standard? Or guideline?).tw.
	48	problem?.tw.
	49	bias?.tw.
	50	ethic*.tw.
	51	consequence?.tw.
	52	critique?.tw.
	53	insecurity.tw.
	54	implications.tw.
	55	peril?.tw.
	56	impact?.tw.
	57	or/ 41-56
		16 and 40 and 57

5.2. Study Selection

Study selection and coding were done using the DistillerSR systematic review software (Evidence Partners, 2022). Using the *a priori* eligibility criteria, we developed questionnaires for selecting citations during discrete title, abstract, and full text review stages. Two reviewers independently selected studies during each screening stage.

Regular meetings to discuss rating discrepancies and to compare working definitions were held during the review of the first 1,000 references in the title screening stage and for the first 100 references during the abstract screening stage. Any conflicts during the title and abstract screening stages were included in the full text review. In the full text screening stage, daily meetings were held during the review of the first 20 references to discuss rating discrepancies and to improve working definitions of terms. Rating discrepancies were resolved by discussion, and in five cases, by using a third adjudicator.

5.2.1. Data Collection Process

For included studies, we extracted details on study characteristics (year of publication, countries of all authors, author organization types), population characteristics (type of humanitarian crisis), intervention characteristics (purpose of data processing, technologies described), and outcomes (specific ethical issues identified, whether studies used real-world examples to identify issues). Author organization types were coded for all listed affiliations, while author country was extracted only from the first-listed affiliation. For each country, we additionally tabulated the geographic region and income level, using the 2020 World Bank classification scheme (World Bank, 2020).

The data extraction form was created in the DistillerSR software. It was then piloted based on a random sample of 10 included studies and modified based on discussions and

feedback from the two reviewers. As per the study protocol, since the number of included citations was greater than 30, data extraction was done by one reviewer and verified by another. The data extraction form included several pre-coded ethical issues, but additional emergent issues could be entered qualitatively in text format.

5.2.2. Synthesis

We summarized results quantitatively (using frequencies) and qualitatively (using descriptive analytics). We analyzed and coded the ethical issues related to data processing that were entered in text form using SPSS 25. Specific issues described by authors could be assigned to one or more categories of ethical issues. Issue codes were updated iteratively and recursively by creating new codes based on new observations and through constant retrospective reviews of previously collected data. In some cases, rarely-mentioned codes were also merged retrospectively to limit the size of the final list of issues. The ethical issues mentioned in each study were then grouped into the ethical value categories of autonomy, beneficence, non-maleficence, and justice, based on which category was deemed to be the affected most.

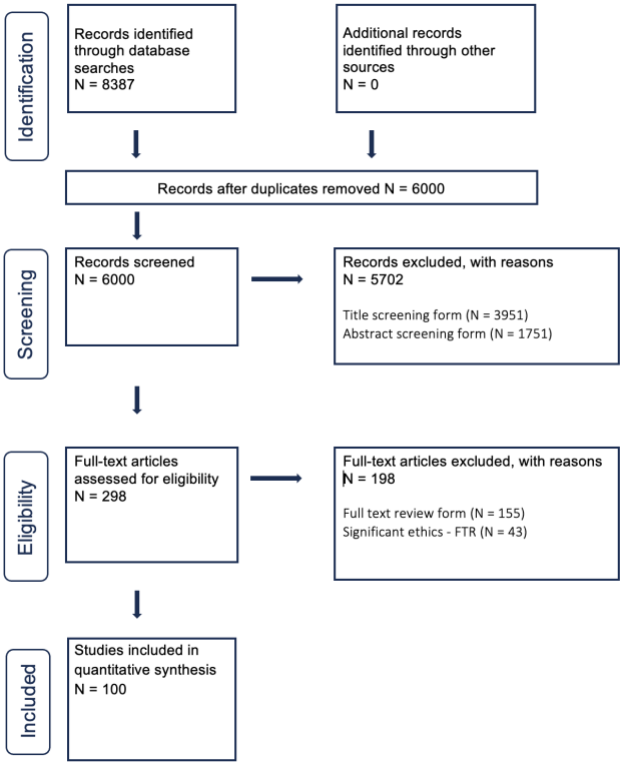
5.3. Results

5.3.1. Literature search

The database literature search returned 8,387 citations (see Figure 7). After removing duplicates, 6,000 were included for screening. 3,951 were excluded during the title screening stage and 1,751 during abstract screening. After reviewing full texts of 298 potentially relevant studies, 198 were excluded. As a result, 100 were included in this scoping review (full list of citations listed in Chapter 5 - Appendix 4).

Figure 7

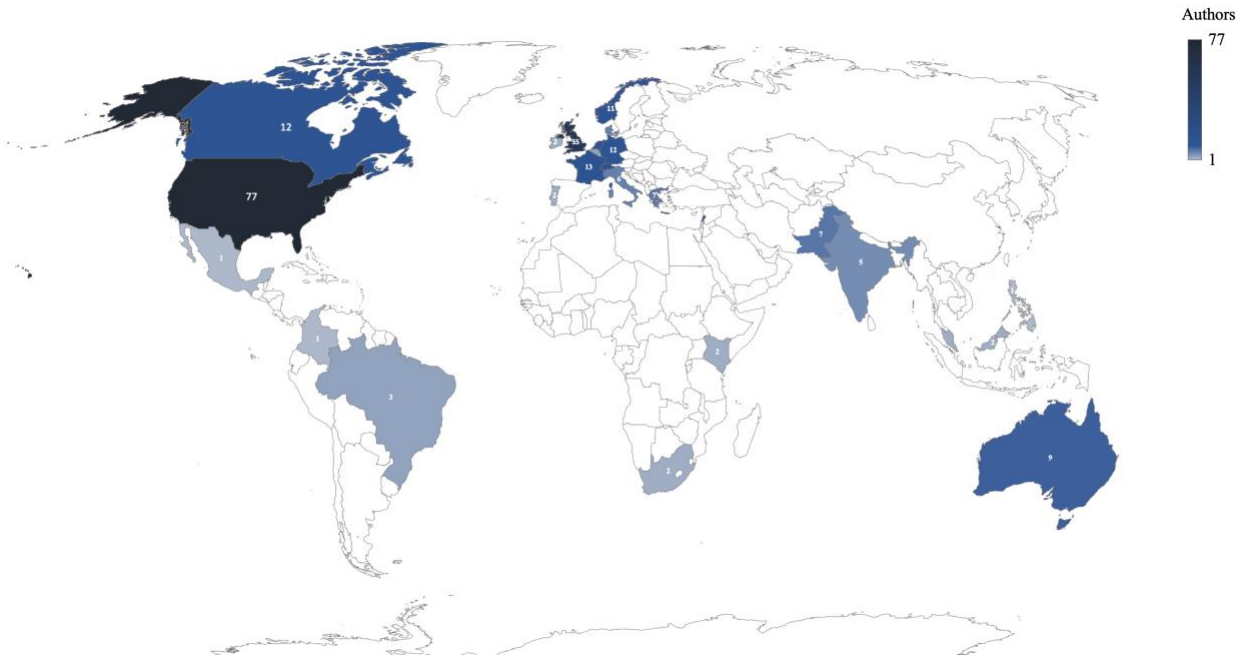
Study Flow Detailing the Flow of Information About Studies During All Stages of the Scoping Review



5.3.2. Study Characteristics

Figure 8

Map Showing the Number of Authors per Country of Primary Affiliation



The included 100 studies were published between 2010 and 2019, as shown in Table 13. The majority (n = 73) were published after 2015, and the most common publication year was 2019 (n = 28). Most were written by authors based in Europe and Central Asia (n = 56) and North America (n = 39), while only a small number of studies included authors from East Asia & Pacific (n = 8), South Asia (n = 5), Sub-Saharan Africa (n = 4), Middle East and North Africa (n = 3), and Latin America and the Caribbean (n = 2), as shown in Figure 8. 32 studies included an author from the United States while about one quarter (n = 27) included an author from the United Kingdom. Overall, 93 studies included at least one author from a high-income country while a smaller number included at least one author from an upper middle-income country (n = 8) or lower middle-income country (n = 8). No study included an author from a low-income country. Similarly, no study included an author from China. The vast majority (n = 92) of studies

included at least one author from an academic institution, while only 7 studies included at least one author affiliated with a humanitarian organization.

Table 13

Study Characteristics

(N=100)		**Count**
Year of publication	2010	2
	2012	2
	2013	5
	2014	9
	2015	9
	2016	20
	2017	12
	2018	13
	2019	28
Region represented by authors	Europe & Central Asia	56
	North America	39
	East Asia & Pacific	8
	South Asia	5
	Sub-Saharan Africa	4
	Middle East & North Africa	3
	Latin America & Caribbean	2
Country income level based on author location	High income	93
	Upper middle-income	8
	Lower middle-income	8
	Low-income	0
Parent organization type based on author affiliation	Academic	92
	For-profit	12
	Non-profit	10
	Humanitarian	7
	Think tank	2
	Government	1

5.3.3. Type of Humanitarian Crisis

Similar numbers of studies focused on or included examples of natural disasters and armed conflict (n = 37 and n = 36, respectively), as shown in Table 14. Of the 100 studies

selected, 32 discussed people displaced by a humanitarian crisis, whereas 19 focused on large public health emergencies. Twenty studies were general in nature and only discussed the fields of humanitarian assistance, emergency management, or disaster response without providing specific examples.

Table 14

Types of Humanitarian Crises Discussed

(N=100)		**Count**
Type of humanitarian crisis	Natural disaster	37
	Armed conflict	36
	People displaced by a humanitarian crisis	32
	Large public health emergency	19
	Not specified	20

5.3.4. Purpose of Data Processing

While most studies reported more than one purpose, the most common data processing purpose was conducting assessments (n = 36), such as needs assessments or damage surveys (see Table 15). Twenty-four studies examined different forms of case management (e.g., refugee registrations), while 20 discussed handling of medical or public health data. Twenty-three did not specify any reasons for data processing but instead discussed in theoretical terms the use of information and communication technologies or data processing in humanitarian assistance.

5.3.5. Technologies Described

The most commonly described technologies used for data processing were social media (discussed by 48 studies), crowdsourcing (n = 45), various forms of mapping and other forms of geographic information systems (GIS; n = 43), whereas nearly one in three studies focused on big data (n = 30) or private messaging services (n = 30).

One quarter (n = 25) discussed various forms of tools and technologies related to artificial intelligence (AI), including the use of algorithms and machine learning. The capture of satellite images for humanitarian assistance and the collection of biometrics (typically, fingerprint or iris scans) were discussed by 23 and 22 studies, respectively. Of the 22 studies, 19 discussing biometrics were published after 2015. Other technologies cited are shown in Table 15. Five studies did not discuss any specific technologies used for data processing.

Table 15

Data Processing Purposes and Technologies Described by Studies

(N=100)		**Count**
Purposes of data processing	Assessment (of needs, damage, etc.)	36
	Registration/case management	24
	Medical care or public health	20
	Forecasting/modeling/early warning	19
	Delivery of assistance	12
	Accountability (complaints, feedback collection, etc.)	9
	Cash transfer	8
	Human rights violations	8
	Logistics	8
	Search and rescue	8
	Other	18
	Not specified	23
Specific technologies described	Social media	48
	Crowdsourcing	45
	Mapping/GIS	43
	Big data	30
	SMS or private messaging software	30
	AI/algorithms/machine learning	25
	Satellite imagery	23
	Biometrics	22
	Information systems	22
	UAV	18
	Medical data	12
	Cash distribution	11

	Call data records	9
	Data storage	5
	Blockchain / distributed ledger technology	4
	Computer-assisted personal interviewing	3
	Not specified	5

5.3.6. Ethical Issues Identified

As shown in Table 16, we identified 22 ethical issues in the studies under investigation, which were grouped according to the four previously identified bioethical values categories. Eight issues were attributed to the ethical value category of autonomy, six to beneficence, seven to non-maleficence, and five to justice. On average, studies cited seven different ethical issues each (M = 6.97, SD = 3.4), ranging from more than two for non-maleficence issues (M = 2.44, SD = 1.3) to less than one for justice issues (M = 0.98, SD = 0.77; see Table 17). The vast majority of studies mentioned issues related to non-maleficence (n = 93) and beneficence (n = 89). Slightly fewer studies discussed issues concerning justice (n = 73) and autonomy (n = 71).

Table 16

Ethical Issues Identified

(N = 100)		**Count**
Autonomy	Lack of consent: Data are collected without informed consent	52
	Data agency: People do not have the right to control, access, or delete their data	31
	Lack of respect: People/communities are not treated with respect	30
	Autonomy: Unwillingness to share data does not lead to disadvantages (e.g., exclusion from assistance or protection	26
	Participation: People/communities are not involved in decisions to use of new/experimental technologies for collecting data	18
	Undisclosed use: Data may be used beyond purposes for which they were collected	8
	Lack of group agency: Processed information is not available to affected communities	8
	Any issue related to Autonomy	*71*
Beneficence	Unreliability: Processed data are inaccurate and do not sufficiently reflect reality to inform assistance	61
	Dependence: Data is processed with the assistance of a political, economic, or military entity	48
	Lack of action: Processed data is not utilized to inform assistance to the affected person/community	40
	Non-neutrality: Data is processed in a way that benefits or appears to benefit one side of the conflict over the other	29
	Ineffective or inefficient: Not producing expected result, unmet expectations	4
	Any issue related to Beneficence	*89*
Non-maleficence	Privacy: Personal/sensitive data is shared with third parties	74
	Harm: People may suffer physical or psychological harm as a result of data processing	68
	Data security: Personal/sensitive data is not protected against malicious actors	42
	Power imbalance: Data processing reinforces or worsens a lack of power of affected people	34
	Excess: More data was collected than necessary	17
	Redress/rectification: People do not have the ability to correct wrong information about them or receive compensation	9
	Any issue related to Non-maleficence	93
Justice	Bias: Data is processed in a way that may result in aid being distributed disproportionately to people's actual needs	63
	Lack of accountability: Endangering (or not protecting) rights; absolving responsibility	16
	Barriers to inclusion due to technology choices during data collection	14
	Unfair distribution of risks and benefits	5
	Any issue related to Justice	*73*

The most frequently cited ethical issue categorized under the value category of autonomy was data being collected without sufficient informed consent (n = 52). For example, Shoemaker et. al. (2019) found through interviews with refugees are frequently being asked by humanitarian organizations to provide personal information that the respondents considered intrusive, without being offered a justification on why this was relevant.

Within the value category of beneficence, the ethical issue most frequently mentioned by studies was processed data being inaccurate and not sufficiently reflecting reality to inform assistance (n = 61). This is illustrated by Paul & Sosale (2020), who cite the challenges of using social media as a basis to inform humanitarian assistance. In an example the authors cite, the same information was re-posted multiple times by well-meaning users, making it difficult for emergency responders after a severe flooding event to identify new information that might require a team to be dispatched.

Falling under the value category of non-maleficence, the most-cited ethical issue (n = 74) were privacy concerns in cases where personal or sensitive data may be shared with third parties. For example, Hayes & Kelly (2018) discuss how personal requests for help that are aggregated by a crowdsourcing platform such as Ushahidi can make personal information publicly available, including to bad actors trying to exploit vulnerable people.

The most frequently mentioned ethical issue categorized under the justice value category was biased data processing leading to (dis)advantaging people disproportionate to their humanitarian needs (n = 63). This issue, which often relates to different forms of sampling problems that could endanger the impartial distribution of aid, has become more pressing as more organizations turn to “big data” solutions for informing humanitarian assistance without properly understanding their limitations (Latif et al., 2019).

Table 17

Number of Ethical Issues Cited by Ethical Value Category

(N = 100)	**mean, SD (min to max)**
All ethical value categories	*6.97, 3.4 (1 to 15)*
Autonomy	1.73, 1.59 (0 to 7)
Beneficence	1.82, 1.16 (0 to 5)
Non-maleficence	2.44, 1.3 (0 to 6)
Justice	0.98, 0.77 (0 to 3)

5.3.7. Information Sources for Ethical Issues

As shown in Table 18, slightly over half of studies (n = 52) cited at least one real-world example of an ethical issue, usually based on anecdotal information found in news reports or other published literature (see, e.g., Comes et al., 2019; Haworth et al., 2018). Fully 29 studies included ethical issues that were raised by interviews or other kinds of consultations with experts. Examples here include Shoemaker et. al. (2019), who conducted qualitative interviews with 198 refugees in Lebanon, Jordan, and Uganda, as well as Vannini et al. (2020), who interviewed nine representatives from organizations assisting transborder migrants in the United States. Four studies included a systematic review of the literature (Bempong et al., 2019; Cinnamon et al., 2016; Mesmar et al., 2016; Watson & Rodrigues, 2018).

Table 18

Information Sources of Ethical Issues

(N = 100)	**Count**
Specific instances of ethical issue (rooted in real-life experience)	52
Ethical concerns raised in interviews/expert consultations	29
Systematic review of the literature	4

5.3.8. Key Results for Studies Discussing AI

Of the 25 studies that discuss the use of AI, all were published since 2014, with about half (n = 13) published after 2017 (see Chapter 5 - Appendix 3 for all figures related to the AI-related studies). The most common type of humanitarian crisis discussed in the 25 studies was natural disaster (n = 9), followed by armed conflict and large public health emergencies (n = 7, respectively). The most common purposes for data processing were related to assessments (n = 9) as well as handling medical or public health data (n = 7). The majority of the 25 studies related to AI also discussed big data (n = 15), social media (n = 15), and GIS (n = 13). The vast majority mentioned ethical implications related to privacy (n = 23) and the risk of physical or psychological harm (n = 22). Eighteen studies related to data being processed in a way that may result in aid being distributed disproportionately with regard to people's actual needs.

5.4. Discussion

The aim of this review was to map the range of ethical issues that have been raised in the academic literature regarding data processing of people affected by humanitarian crises. This review identified 22 such ethical issues. Issues related to the value category of non-maleficence were brought up by the vast majority of studies (n = 93), which dovetails with a strong trend in the recent literature focusing on the imperative of "do no harm" in humanitarian assistance (Anderson, 1999; Sandvik & Raymond, 2017; WHS, 2016). The risk of increasing harm (whether physical or psychological) as a result of data processing was mentioned by a high number of studies (n = 68).

Privacy concerns were cited by 74 studies—far more commonly than all other issues—reflecting an increased awareness of this issue over the last years across organizations and the

media. Among studies discussing AI as a data processing technology, privacy concerns were even more prevalent, with 23 out of 25 (92%) mentioning this issue. This points to a significant worry across the humanitarian sector about the many ways in which personal data from affected people is being processed in a manner that may endanger their right to privacy, which is enshrined in the 1948 Universal Declaration of Human Rights (UN General Assembly, 1948). Studies discuss a wide gamut of how personal privacy can be violated, including accidental or intentional sharing with third parties beyond what the affected person had agreed to during personal interviews, or if at all. Even in cases where informed consent was given, interviewees in vulnerable situations—or who lack understanding of sophisticated data management, access and processing—may not understand all the potential ways in how their personal information may be used, stored and accessed. Collecting and processing personal data from social media, by unmanned aerial vehicles (UAV), or public records (often under the "big data" category) that lack explicit consent are particularly problematic. Although the protection of privacy can be understood an essential right to safeguard human dignity (Floridi, 2016), more studies and initiatives in humanitarian assistance need to resolve the apparent conflict between the duty to protect privacy and the urgent duty to assist and protect those in danger (Hayes & Kelly, 2018).

Many studies pointed out that organizations frequently collect much more data than they need (n = 17) or are able to absorb (n = 40). We consider the former as a potential for harm, as any excessive information increases the risks associated with data leaks and privacy violations. Data collected but that is not used can be regarded as related to the ethical value of beneficence, which implies that all information collected should have a concrete purpose related to informing humanitarian assistance. But even for data that were used for the intended purpose, a majority (n

= 61) of studies discussed that it may be too unreliable or inaccurate to adequately inform assistance programming.

Related to the ethical value category of autonomy, a strong theme emerged regarding insufficient consent mechanisms. About half the studies (n = 52) mentioned that informed consent was either not provided by the affected population or was given without a full picture of how data would be processed or used. Eight studies cited that data might be used for reasons other than the original purpose for which consent may have been obtained. Related to the ethical value category of autonomy, about one in four papers (n = 26) mentioned that a refusal to provide information could lead to being excluded from receiving assistance. This issue is illustrated by Shoemaker (2019) who documented how refugees felt that they lacked a choice on whether or not to provide personal information to UNHCR as their ability to access assistance depended on it. Detailed guidance has been created by the International Rescue Committee on how to obtain proper consent, whereas the International Committee of the Red Cross has published the legal basis for situations when data processing is permissible—even when consent cannot be assumed or obtained (International Committee of the Red Cross, 2020; International Rescue Committee, 2018). However, more work is clearly needed to train humanitarian professionals in these practices, and to monitor for better compliance with best consent practices as well as other minimal ethical guidelines. Existing guidance also needs to be updated to ensure the protection of private, personal and demographically identifiable information that extends to population groups rather than individuals (Raymond, 2017).

Directly related to the value category of justice, a majority (n = 63) of studies were concerned about data being processed in a way that may result in aid being distributed disproportionately with regard to people's actual needs. This finding directly mirrors the

importance of the humanitarian impartiality value category which refers to providing assistance solely based on need and regardless of personal preferences or discriminatory factors (Pictet, 1979).

A cross-cutting issue was the potential of data processing to exacerbate power imbalances (mentioned by 34 studies), often due to an exclusion from data collection, given the unequal access to certain technologies (n = 14). In many cases, data processing was found to diminish the perceived neutrality of humanitarian organizations (n = 29) as data could be processed in a way that might benefit one side of the conflict over the other. Concerningly, about half (n = 48) of the studies found that humanitarian data processing might be overly dependent on potentially biased external entities (such as commercial entities, militaries, or foreign governments). This could be increasingly problematic for humanitarian organizations for multiple operational and ethical reasons, but particularly in conflict environments where the perception of independence is widely considered to be an essential humanitarian value category.

Another theme identified across many studies was that data processing did not follow the principles of Accountability to Affected People (AAP) (Global Health Cluster, 2017; International Committee of the Red Cross, 2019), which manifested in various ways across several of the four bioethical value categories. For the value category of autonomy, 18 studies remarked that affected communities were not involved in decisions on whether to use experimental technologies, whereas a smaller number (n = 8) commented that processed information was not being made available to communities to allow for better group agency. Related to the values category of non-maleficence, nine papers discussed people's inability to rectify inaccurate information about them or receive any form of compensation. Finally, related to the value category of justice, one in six (n = 16) of studies found that data processing lacked

accountability in terms of humanitarian organizations' obligation to protect rights—or even pointed to ways that they may be violating these rights themselves. The Signal Code (Greenwood et al., 2017), first published by the Harvard Humanitarian Initiative in 2017, considers data agency and redress/rectification as crucial rights and proposes specific actions to safeguard them in practice. We propose extending this list to always include affected communities when sharing collected data and involving them in decisions over experimental technologies.

About half of studies (n = 52) cited ethical issues that were rooted in real-life experiences whereas almost one third (n = 29) contained issues based on qualitative interviews or expert consultations. This signals that ethical issues have moved from theoretical concerns to actual incidents. However, it also reflects the large and diverse array of ethical issues that are emerging in connection with data processing in humanitarian crises which may first manifest as theoretical concerns before being validated as potentially negative consequences that can and do occur in real life.

The results from this study show a wide array of ethical issues that should be addressed when processing data in a humanitarian context. However, to our knowledge, to date no humanitarian data protection guidance is sufficiently comprehensive to provide practical guidance for all concerns identified in the literature. More work may therefore be needed to expand existing guidelines and to ensure that future updates are also informed by systematic reviews of newly published studies. Moreover, there is a need to improve humanitarian organizations' accountability to affected populations, including the need to proactively prevent harm, and monitor the potential for causing harm and to limit risks. Finally, organizations involved in providing humanitarian assistance should review internal processes for training staff

and create methods for verification that ensure appropriate minimal ethical standards are being met.

Geographically, publications were disproportionately from authors in high-income countries, primarily in Europe and North America, demonstrating a high level of interest in countries that have been the traditional source of most humanitarian funding but also of most technological innovation. Conversely, the small number of authors from lower-middle income countries and the complete absence of any authors from low-income countries highlights the lack of published perspectives from countries most affected by humanitarian crises. The lack of any study with an author from China may reflect that the large body of disaster related studies from Chinese authors published in English primarily discuss the response to domestic rather than foreign crises, or that ethical issues explored in this review may be more explored in Chinese language publications. People living in affected countries make up the vast majority of humanitarian organizations' staff, which could be a potential boon to a more diverse authorship on this subject. However, given the very small number of studies with authors from a humanitarian organization, more efforts need to be made by publishers to invite and support submissions from humanitarian professionals.

The number of studies discussing natural disasters (n = 37) was about the same as the number discussing armed conflict (n = 36), even though by the end of 2020, 87% of displacement was caused by conflict (IDMC, 2021). This disproportionate focus may be due to disasters generating a higher level of media attention, as well as interest among technology enthusiasts, volunteers, and private companies—a trend identified by several studies (Burns, 2019; Duffield, 2016; Lev Aretz, 2019; Taylor, 2016). Likewise, empirical research in conflict settings is far more difficult given the inherent security risks, which in turn limits the

development of theories and academic discourse that rely on data from the field. More research will be needed in the future focusing on ethical issues that are unique to conflict settings, as data processing without appropriate consideration of ethical issues in these settings arguably has the potential to cause far greater harm.

Our results show a significant focus on both internally and externally displaced populations, particularly those trying to reach Europe or the United States. Likewise, the Ebola virus disease outbreak in West Africa in 2014-2016 was the focus of major international containment response because of the perceived threat of pandemic spread beyond the region, and the subject of a large number of well-funded studies. The increased focus on assistance to displaced populations could be due to the intense media coverage, whereas technological experimentation during the Ebola response was taken to new levels in areas such as processing of call data records without explicit consent (Sean McDonald, 2016).

A significant number of studies discussed ethical issues without going into detail about the particular context: 20 studies discussed humanitarian assistance or crisis response in general, while 23 did not specify a data processing purpose. Given the lack of widely shared understandings of what constitutes terms such as “humanitarian community” or “information and communication technologies”, we recommend that even theoretical papers provide sufficient definitions and examples.

We found that studies most commonly discussed activities involving the initial collection of data from affected populations, including assessments, registrations, and health interventions. To some extent, this reflects that a large number of studies investigated the use of crowdsourcing and social media to gain an understanding of a particular humanitarian crisis (see below). It may also be a reflection of the increasing emphasis that humanitarian organizations and their donors

have been placing in recent years on establishing an “evidence base” before rolling out assistance programs (Blanchet et al., 2017; K. T. Pham et al., 2017). More research is needed to investigate the link between the potential increase of ethical risk and the push for collecting more needs assessment data.

Studies discussing social media (n = 48), crowdsourcing (n = 45), and mapping (n = 43) dominated, often due to the perceived lack of good ground-validated data in humanitarian assistance. There were many use cases of social media, but the most-discussed application was mining public Twitter posts for clues on potential population needs. We also found that many studies focus on the potential use of other “new” technologies, especially if they can be used remotely to assess needs (e.g., satellite imagery, unmanned aerial vehicles, call data records). Crowdsourcing, a method of obtaining information from the general public (Martin-Shields, 2013), was discussed by almost half the studies. Many studies traced their enthusiasm for—or criticism of—crowdsourcing to the creation of the Ushahidi platform (mentioned by 25 studies) in 2007. Similarly, the emergence of digital platform based volunteer networks since the 2010 Haiti earthquake (Meier, 2014; Phillips, 2018) can partially explain the large number of studies referencing these tools.

Surprisingly, only three studies mentioned computer-assisted personal interviewing (CAPI) tools such as KoboToolbox which has been adopted by a broad range of international and national humanitarian agencies as the tool of choice for needs assessments (OCHA, 2015; Sapkota & Siddiqi, 2019). Similarly, use of spreadsheets were only mentioned by one study as a cause for ethical concern, despite being the main data storage and sharing mechanism of choice for many humanitarian organizations (Madon & Schoemaker, 2019). Such low-tech data processing means are addressed in recent guidelines, for example, giving guidance on how to

remove sensitive data before sharing Excel files with others (OCHA, 2019). However, more research is needed on current practices and ethical risks associated with these commonly used technologies.

The ethical issues associated with biometrics were discussed by a significant number of studies, particularly for the registration of refugees and other migrants by organizations such as the UN High Commissioner for Refugees, UNHCR (see, for example, Jacobsen, 2015). In 2015, such concerns led Oxfam, one of the largest international humanitarian NGOs, to put a moratorium on its use of biometrics in order to assess potential risks (The Engine Room and Oxfam, 2018). In 2021, this in turn resulted in the creation of a policy intended to ensure that the technology is used ethically within Oxfam's operations (Eaton-Lee & Shaughnessy, 2021).

Finally, ethics related to AI and similar technologies were discussed significantly more frequently after 2017. This seems to correlate with the growing presence and desire to analyze "big data" resources, in order to learn more about the needs and sentiments of the affected population. For example, big data and medical data were mentioned twice as frequently by studies that discussed AI; call data records were cited 78% more often. More theoretical and empirical research is needed to understand the potential issues that come with applying these rapidly evolving technologies to humanitarian assistance.

5.4.1. Limitations of the Scoping Review

This study is limited to literature published since 2010 and before January 2020, and it excludes work from non-peer-reviewed sources. As mentioned above, identifying all relevant studies was a significant challenge due to the lack of a shared nomenclature across disciplines for humanitarian assistance, ethical issues, and data processing. As a result, potentially relevant articles that met the inclusion criteria may have been missed. Nonetheless, we believe that our

search strategy represents the most comprehensive and inclusive set of keywords to capture studies in the diverse field of humanitarian assistance to date.

As suggested by the Arksey and O'Malley framework, a consultation exercise with humanitarian and ethics experts will be organized to present our results, aid knowledge translation, ensure that the results from this study are relevant, and frame a future research agenda. The results of this consultation will be published separately.

5.5. Conclusion

From 2010 to 2019, the number of peer-reviewed articles and conference proceedings about the ethical issues related to processing data of people affected by humanitarian crises has grown significantly, particularly after 2015. Published research shows widespread concerns that data processing in humanitarian assistance, without due ethical attention to known harms or potential risks, can cause additional harm, may not provide direct benefit, can show a lack of respect for affected populations' autonomy, and can lead to the unfair distribution of resources, among other concerns. We found that studies containing ethical discussions are often skewed towards investigating natural disaster contexts, as well as the use of technologies that allow the involvement of non-traditional actors, especially by gathering information from social media or crowdsourcing platforms. More research into ethical issues that arise in conflict settings is needed to better investigate the heightened security risks to vulnerable people in the large number of humanitarian crises associated with conflict, war, and social instability.

As some studies in this review have pointed out, data may be collected in many humanitarian crises without due attention to data privacy and the security of the person providing the data, leaving them subject to malfeasant actors. Further, data can be collected or processed in a way that is inaccurate or out-of-date for the purpose for which it is being used, a risk that

increases with methods such as crowdsourcing or analyzing social media posts. Both potential outcomes may pose added safety and security risks for people already vulnerable in a humanitarian crisis. They may also potentially render inaccurate needs assessments to humanitarian actors that may mean material humanitarian assistance delivered is inappropriate to actual needs, thus further increasing the vulnerability of people already at risk.

The ethical issues identified in this review should be used to inform the development of ethical codes of conduct (whether voluntary or mandated by organizations). Further, companies and institutions behind the various technologies—as well as the humanitarian organizations that use them to process data as part of their work—should investigate to what extent these ethical issues are being addressed, and where more needs to be done. However, existing humanitarian data protection guidance and mechanisms are not sufficient to address all concerns identified in the literature and in this study. Likewise, training and accountability mechanisms to monitor the actual harm or potential for causing harm and to limit risks, are insufficient. These guidelines and mechanisms will need to be reviewed, expanded and informed by regular reviews that keep pace with technological change and changes in practice. Further research, especially using empirical methods, is necessary to better identify and understand the type and prevalence of ethical issues in the field.

Finally, more investigations are needed into the appropriate and inappropriate use of commonplace humanitarian tools and data management processes, such as CAPI, spreadsheets, filesharing, or use of online databases. At the same time, case studies of early adoptions of AI should address which ethical considerations were given when using tools that may involve data processing using multiple services and companies globally, in order to inform local decisions.

Such research is urgently needed to create better guidance, training, and auditing methods to support humanitarian organizations to use data processing technologies as ethically as possible.

5.6. References

Anderson, M. B. (1999). *Do no harm: How aid can support peace – Or war*. Lynne Rienner Publishers.

Arksey, H., & O'Malley, L. (2005). Scoping studies: Towards a methodological framework. *International Journal of Social Research Methodology*, *8*(1), 19–32. https://doi.org/10.1080/1364557032000119616

Beauchamp, T. L., & Childress, J. F. (2019). *Principles of Biomedical Ethics* (7th editio). Oxford University Press.

Bempong, N. E., De Castañeda, R. R., Schütte, S., Bolon, I., Keiser, O., Escher, G., & Flahault, A. (2019). Precision Global Health - The case of Ebola: A scoping review. *Journal of Global Health*, *9*(1). https://doi.org/10.7189/jogh.09.010404

Blanchet, K., Ramesh, A., Frison, S., Warren, E., Hossain, M., Smith, J., Knight, A., Post, N., Lewis, C., Woodward, A., Dahab, M., Ruby, A., Sistenich, V., Pantuliano, S., & Roberts, B. (2017). Evidence on public health interventions in humanitarian crises. *The Lancet*, *390*(10109), 2287–2296. https://doi.org/10.1016/S0140-6736(16)30768-1

Broussard, G., Rubenstein, L. S., Robinson, C., Maziak, W., Gilbert, S. Z., & DeCamp, M. (2019). Challenges to ethical obligations and humanitarian principles in conflict settings: a systematic review. *Journal of International Humanitarian Action*, *4*(1). https://doi.org/10.1186/s41018-019-0063-x

Burns, R. (2019). New Frontiers of Philanthro-capitalism: Digital Technologies and Humanitarianism. *Antipode*, *51*(4), 1101–1122. https://doi.org/10.1111/anti.12534

Cawthorne, D., & Wynsberghe, A. R. Van. (2019). From HealthDrone to FrugalDrone: Value-Sensitive Design of a Blood Sample Transportation Drone. *International Symposium on*

Technology and Society, Proceedings, *2019-Novem*. https://doi.org/10.1109/ISTAS48451.2019.8938019

Cinnamon, J., Jones, S. K., & Adger, W. N. (2016). Evidence and future potential of mobile phone data for disease disaster management. *Geoforum*, *75*, 253–264. https://doi.org/10.1016/j.geoforum.2016.07.019

Comes, T., Meesters, K., & Torjesen, S. (2019). Making sense of crises: the implications of information asymmetries for resilience and social justice in disaster-ridden communities. *Sustainable and Resilient Infrastructure*, *4*(3), 124–136. https://doi.org/10.1080/23789689.2017.1405653

Donini, A., & Maxwell, D. (2013). From face-to-face to face-to-screen: Remote management, effectiveness and accountability of humanitarian action in insecure environments. *International Review of the Red Cross*, *95*(890), 383–413. https://doi.org/10.1017/S1816383114000265

Duffield, M. (2016). The resilience of the ruins: towards a critique of digital humanitarianism. *Resilience*, *4*(3), 147–165. https://doi.org/10.1080/21693293.2016.1153772

Eaton-Lee, J., & Shaughnessy, E. (2021, June 25). *Oxfam's new policy on biometrics explores safe and responsible data practice*. Reliefweb. https://reliefweb.int/report/world/oxfam-s-new-policy-biometrics-explores-safe-and-responsible-data-practice

Elhra, & Humanitarian Health Ethics. (2021). *Ethics for humanitarian innovation: Toolkit*. https://higuide.elrha.org/wp-content/uploads/2021/09/Elrha-Ethics_for_Humanitarian_Innovation-Toolkit-210913.pdf

Evidence Partners. (2022). *DistillerSR*. https://www.evidencepartners.com

Floridi, L. (2016). On Human Dignity as a Foundation for the Right to Privacy. In *Philosophy*

and Technology (Vol. 29, Issue 4). https://doi.org/10.1007/s13347-016-0220-8

Global Health Cluster. (2017). *Operational Guidance on Accountability to Affected Populations (AAP)*. https://www.humanitarianresponse.info/system/files/documents/files/cluster_coordination_reference_module_2015_final.pdf

Greenwood, F., Howarth, C., Poole, D. E., Raymond, N. A., & Scarnecchia, D. P. (2017). *The Signal Code: A Human Rights Approach to Information During Crisis*. Cambridge, MA: Harvard Humanitarian Initiative. https://signalcodeorg.files.wordpress.com/2017/01/signalcode_final7.pdf

Haworth, B. T., Bruce, E., Whittaker, J., & Read, R. (2018). The good, the bad, and the uncertain: Contributions of volunteered geographic information to community disaster resilience. *Frontiers in Earth Science*, *6*(November), 1–15. https://doi.org/10.3389/feart.2018.00183

Hayes, P., & Kelly, S. (2018). Distributed morality, privacy, and social media in natural disaster response. *Technology in Society*, *54*, 155. http://ezproxy.library.yorku.ca/login?url=https://search.proquest.com/docview/2113599412?accountid=15182

Hilhorst, D., & Schmiemann, N. (2002). Development in Practice Humanitarian principles and organisational culture: Everyday practice in Médecins Sans Frontières-Holland. *Development in Practice*, *12*(3–4), 490–500. https://doi.org/10.1080/0961450220149834

Humanitarian Outcomes. (2021). *Aid Worker Security Database*. https://aidworkersecurity.org/incidents/report

IASC. (1999). *Protection of Internally Displaced Persons*.

https://interagencystandingcommittee.org/system/files/legacy_files/protection_of_internally_displaced_persons_inter_agency_standing_committee_policy_paper_0.pdf

IASC. (2021). *Operational Guidance: Data Responsibility in Humanitarian Action*. Inter-Agency Standing Committee. https://interagencystandingcommittee.org/operational-response/iasc-operational-guidance-data-responsibility-humanitarian-action

IDMC. (2021). *2021 Global Report on Internal Displacement*. Internal Displacement Monitoring Centre. https://www.internal-displacement.org/global-report/grid2021/

IFRC. (1994). *The Code of Conduct for the International Red Cross and Red Crescent Movement and Non Governmental Organisations (NGOs) in Disaster Relief*. http://www.ifrc.org/Global/Publications/disasters/code-of-conduct/code-english.pdf

International Committee of the Red Cross. (2019). *Accountability to Affected People Institutional Framework*. https://www.icrc.org/en/publication/accountability-affected-people-institutional-framework

International Committee of the Red Cross. (2020). *Handbook on Data Protection in Humanitarian Action (2nd Edition)*. https://www.icrc.org/en/data-protection-humanitarian-action-handbook

International Committee of the Red Cross and Privacy International. (2018). *The Humanitarian Metadata Problem: "Doing No Harm" in the Digital Era*. https://privacyinternational.org/sites/default/files/2018-12/The Humanitarian Metadata Problem - Doing No Harm in the Digital Era.pdf

International Rescue Committee. (2018). *Obtaining Meaningful Informed Consent*. https://www.principlesinpractice.info/help-library/irc-research-toolkit-obtaining-meaningful-informed-consent

Jacobsen, K. L. (2015). Experimentation in humanitarian locations: UNHCR and biometric registration of Afghan refugees. *Security Dialogue*, *46*(2), 144–164. https://doi.org/10.1177/0967010614552545

Krishnaraj, G., Hunt, M., Schwartz, L., O'Mathúna, D., Pringle, J., Okhowat, A., Kapiriri, L., Brahimi, L., & Yantzi, R. (2021). *Ethics for Humanitarian Innovation: Background Paper*. Elrha. https://higuide.elrha.org/wp-content/uploads/2021/09/Elrha-Ethics_for_Humanitarian_Innovation-Background_Paper-210909.pdf

Latif, S., Qayyum, A., Usama, M., Qadir, J., Zwitter, A., & Shahzad, M. (2019). Caveat Emptor: The Risks of Using Big Data for Human Development. *IEEE Technology & Society Magazine*, *38*(3), 82–90. http://10.0.4.85/MTS.2019.2930273

Lev Aretz, Y. (2019). Data Philanthropy. *SSRN Electronic Journal*, *70*(6). https://doi.org/10.2139/ssrn.3320798

Levac, D., Colquhoun, H., & O'Brien, K. K. (2010). Scoping studies: Advancing the methodology. *Implementation Science*, *5*(1). https://doi.org/10.1186/1748-5908-5-69

Madon, S., & Schoemaker, E. (2019). Reimagining Refugee Identity Systems: A Sociological Approach. *Information and Communication Technologies for Development: Strengthening Southern-Driven Cooperation as a Catalyst for Ict4d, Pt I*.

Makhoul, J., Chehab, R. F., Shaito, Z., & Sibai, A. M. (2018). A scoping review of reporting "Ethical Research Practices" in research conducted among refugees and war-affected populations in the Arab world. *BMC Medical Ethics*, *19*(1), 1–9. https://doi.org/10.1186/s12910-018-0277-2

Martin-Shields, C. (2013). The Technologists Dilemma: Ethical Challenges of Using Crowdsourcing Technology in Conflict and Disaster-Affected Regions. *Georgetown*

Journal of International Affairs, *14*(2), 157–163. http://www.jstor.com/stable/43134422

Meier, P. (2014). *Digital humanitarians : how big data is changing the face of humanitarian response*. CRC Press.

Mesmar, S., Talhouk, R., Akik, C., Olivier, P., Elhajj, I. H., Elbassuoni, S., Armoush, S., Kalot, J., Balaam, M., Germani, A., & Ghattas, H. (2016). The impact of digital technology on health of populations affected by humanitarian crises: Recent innovations and current gaps. *Journal of Public Health Policy*, *37*(2), S167–S200. https://doi.org/10.1057/s41271-016-0040-1

MSF. (n.d.). *Independent, impartial, neutral*. Retrieved January 4, 2021, from https://www.doctorswithoutborders.org/who-we-are/principles/independence

OCHA. (n.d.). *Financial Tracking Service*. Retrieved February 1, 2020, from https://fts.unocha.org/

OCHA. (2012). *Humanitarian Principles*. UN Office for the Coordination of Humanitarian Affairs. https://www.unocha.org/sites/dms/Documents/OOM-humanitarianprinciples_eng_June12.pdf

OCHA. (2015). World Humanitarian Data and Trends 2015. In *United Nations Office for the Coordination of Humanitarian Affairs*. https://www.unocha.org/sites/dms/Documents/WHDT2015_2Dec.pdf

OCHA. (2019). *Data responsibility guidelines*. UN Office for the Coordination of Humanitarian Affairs. https://centre.humdata.org/wp-content/uploads/2019/03/OCHA-DR-Guidelines-working-draft-032019.pdf

OCHA. (2021). *Global Humanitarian Overview 2021*. Office for the Coordination of Humanitarian Affairs. https://doi.org/10.18356/9789214030751

Orbinski, J. (1994). Health Impact of Rights Violations in Conf1ict Zones and the Humanitarian Imperative: Rwanda-A Case Study. *Refuge: Canada's Journal on Refugees*, *14*(5), 5–8. https://doi.org/10.25071/1920-7336.21827

Pal, N. E., Eckenwiler, L., Hyppolite, S.-R., Pringle, J., Chung, R., & Hunt, M. (2019). Ethical considerations for closing humanitarian projects: a scoping review. *Journal of International Humanitarian Action*, *4*(1). https://doi.org/10.1186/s41018-019-0064-9

Parker, B. (2019, February 5). *New UN deal with data mining firm Palantir raises protection concerns*. The New Humanitarian. https://www.irinnews.org/news/2019/02/05/un-palantir-deal-data-mining-protection-concerns-wfp

Paul, S., & Sosale, S. (2020). Witnessing a Disaster: Public Use of Digital Technologies in the 2015 South Indian Floods. *Digital Journalism*, *8*(1), 15–31. https://doi.org/10.1080/21670811.2019.1636693

Perakslis, E. D. (2018). Using digital health to enable ethical health research in conflict and other humanitarian settings. *Conflict and Health*, *12*(1), 23. https://doi.org/10.1186/s13031-018-0163-z

Peters, M. D. J., Godfrey, C. M., Khalil, H., McInerney, P., Parker, D., & Soares, C. B. (2015). Guidance for conducting systematic scoping reviews. *International Journal of Evidence-Based Healthcare*, *13*(3), 141–146. https://doi.org/10.1097/XEB.0000000000000050

Peters, M. D. J., Godfrey, C., Mcinerney, P., Baldini Soares, C., Khalil, H., & Parker, D. (2017). Chapter 11: Scoping reviews. In E. Aromataris & Z. Munn (Eds.), *Joana Briggs Institute Reviewer's Manual* (Issue September, pp. 6–24). https://reviewersmanual.joannabriggs.org/

Pham, K. T., Sattigeri, P., Dhurandhar, A., Jacob, A. C., Vukovic, M., Chataigner, P., Freire, J., Mojsilovic, A., & Varshney, K. R. (2017). Real-time understanding of humanitarian crises

via targeted information retrieval. *IBM Journal of Research and Development*, *61*(6), 7:1-7:12. https://doi.org/10.1147/JRD.2017.2722799

Pham, P. N., & Vinck, P. (2012). Technology fusion and their implications for conflict early warning systems, public health, and human rights. *Health and Human Rights*, *14*(2), 106–117. http://www.jstor.org/stable/healhumarigh.14.2.106

Phillips, J. (2018). Risk in a digital age: understanding risk in virtual networks through digital response networks (DRNs). *International Development Planning Review*, *40*(3), 239–272. https://doi.org/10.3828/idpr.2018.18

Pictet, J. (1979). *The fundamental principles of the Red Cross*. http://www.ifrc.org/PageFiles/40669/Pictet Commentary.pdf

Raymond, N. A. (2017). Beyond "Do No Harm" and Individual Consent: Reckoning with the Emerging Ethical Challenges of Civil Society's Use of Data. In B. Taylor, L and Floridi, L and VanDerSloot (Ed.), *Group Privacy: New Challenges of Data Technologies* (Vol. 126, pp. 67–82). Springer. https://doi.org/10.1007/978-3-319-46608-8_4

Raymond, N. A., Al Achkar, Z., Verhulst, S., & Berens, J. (2016). Building data responsibility into humanitarian action. *OCHA Policy and Studies Series*, *018*, 18. https://docs.unocha.org/sites/dms/Documents/TB18_Data Responsibility_Online.pdf

Raymond, N. A., & Card, B. (2015). *Applying Humanitarian Principles to Current Uses of Information Communication Technologies: Gaps in Doctrine and Challenges to Practice*. Harvard Humanitarian Initiative. https://hhi.harvard.edu/sites/default/files/publications/signal_program_humanitarian_principles_white_paper.pdf

Sandvik, K. B., Gabrielsen Jumbert, M., Karlsrud, J., & Kaufmann, M. (2014). Humanitarian

technology: a critical research agenda. *International Review of the Red Cross*, *96*(893), 219–242. https://doi.org/10.1017/s1816383114000344

Sandvik, K. B., & Raymond, N. A. (2017). Beyond the Protective Effect: Towards a Theory of Harm for Information Communication Technologies in Mass Atrocity Response. *Genocide Studies and Prevention: An International Journal*, *11*(1), 9–24. https://doi.org/10.5038/1911-9933.11.1.1454

Sapkota, P. P., & Siddiqi, K. (2019). Is Ubiquitous Technology for Needs Data Management a Game Changer in Humanitarian Arena? *International Journal of Information Systems for Crisis Response and Management*, *11*(1), 83–97. https://doi.org/10.4018/IJISCRAM.2019010105

Schofield, G., Dittborn, M., Selman, L. E., & Huxtable, R. (2021). Defining ethical challenge(s) in healthcare research: a rapid review. *BMC Medical Ethics*, *22*(1), 1–17. https://doi.org/10.1186/s12910-021-00700-9

Sean McDonald. (2016). *Ebola: A Big Data Disaster* (2016.01; CIS Papers). https://cis-india.org/papers/ebola-a-big-data-disaster

Shoemaker, E., Kristinsdottir, G. S., Ahuja, T., Baslan, D., Pon, B., Currion, P., Gumisizira, P., & Dell, N. (2019). Identity at the margins: Examining refugee experiences with digital identity systems in Lebanon, Jordan, and Uganda. *COMPASS 2019 - Proceedings of the 2019 Conference on Computing and Sustainable Societies*, 206–217. https://doi.org/10.1145/3314344.3332486

Slim, H. (2015). *Humanitarian Ethics: A Guide to the Morality of Aid in War and Disaster*. Oxford University Press.

Sphere Association. (2018). *The Sphere Handbook: Humanitarian Charter and Minimum*

Standards in Humanitarian Response (4th editio). Practical Action Publishing. https://doi.org/ISBN 978-1-908176-00-4

Tansey, C. M., Pringle, J., Davé, A., Boulanger, R., & Hunt, M. (2018). Earthquakes to Floods: A Scoping Review of Health-related Disaster Research in Low- and Middle-income Countries. *PLoS Currents*. https://doi.org/10.1371/currents.dis.57d98a902a326361d88d54521e68b016

Taylor, L. (2016). The ethics of big data as a public good: which public? Whose good? *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *374*(2083), 20160126. https://doi.org/10.1098/rsta.2016.0126

Tegmark, M. (2017). *Life 3.0 : being human in the age of artificial intelligence*. Alfred A. Knopf.

The Engine Room and Oxfam. (2018). *Biometrics in the Humanitarian Sector*. https://www.theengineroom.org/wp-content/uploads/2018/03/Engine-Room-Oxfam-Biometrics-Review.pdf

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D. J., Horsley, T., Weeks, L., Hempel, S., Akl, E. A., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., Wilson, M. G., Garritty, C., … Straus, S. E. (2018). PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. *Annals of Internal Medicine*, *169*(7), 467. https://doi.org/10.7326/M18-0850

Tricco, A. C., Zarin, W., Rios, P., Pham, B., Straus, S. E., & Langlois, E. V. (2016). Barriers, facilitators, strategies and outcomes to engaging policymakers, healthcare managers and policy analysts in knowledge synthesis: a scoping review protocol. *BMJ Open*, *6*(12), e013929. https://doi.org/10.1136/bmjopen-2016-013929

UN General Assembly. (1948). Universal Declaration of Human Rights. In *General Assembly*

resolution. https://www.un.org/en/ga/search/view_doc.asp?symbol=A/RES/217(III)

UN Global Pulse. (2016). *Big Data for Development and Humanitarian Action: Towards Responsible Governance: Global Pulse Privacy Advisory Group Meetings*. http://unglobalpulse.org/sites/default/files/Big_Data_for_Development_and_Humanitarian_Action_Report_Final_0.pdf

UN News Centre. (2015, February 6). *Interview: Senior UN relief official warns no room for complacency until Ebola is eliminated*. https://ebolaresponse.un.org/interview-senior-un-relief-official-warns-no-room-complacency-until-ebola-eliminated

UN OCHA. (2003). *Glossary of Humanitarian Terms*. OCHA. https://www.humanitarianresponse.info/sites/www.humanitarianresponse.info/files/documents/files/ocha glossary.pdf

UNHCR. (2021). *Global Trends: Forced Displacement in 2020*. UNHCR. https://www.unhcr.org/60b638e37/unhcr-global-trends-2020

van Wynsberghe, A., & Comes, T. (2020). Drones in humanitarian contexts, robot ethics, and the human–robot interaction. *Ethics and Information Technology*, *22*(1), 43–53. https://doi.org/10.1007/s10676-019-09514-1

Vannini, S., Gomez, R., Lopez, D., Mora, S., Morrison, J. C., Tanner, J., Youkhana, L., Vergara, G., & del Mar Moreno Tafurt, M. (2020). Humanitarian organizations' information practices: Procedures and privacy concerns for serving the undocumented. *Electronic Journal of Information Systems in Developing Countries*, *86*(1), 1–8. https://doi.org/10.1002/isd2.12109

Wang, N., Christen, M., & Hunt, M. (2021). Ethical Considerations Associated with "Humanitarian Drones": A Scoping Literature Review. *Science and Engineering Ethics*,

27(4), 1–21. https://doi.org/10.1007/s11948-021-00327-4

Watson, H., & Rodrigues, R. (2018). Bringing Privacy into the Fold: Considerations for the Use of Social Media in Crisis Management. *Journal of Contingencies & Crisis Management*, *26*(1), 89–98. http://10.0.4.87/1468-5973.12150

WHO. (2007). *Risk reduction and emergency preparedness*. World Health Organization. https://apps.who.int/iris/bitstream/handle/10665/43736/9789241595896_eng.pdf

WHS. (2016). *The Grand Bargain – A Shared Commitment to Better Serve People in Need*. Istanbul: World Humanitarian Summit. http://reliefweb.int/sites/reliefweb.int/files/resources/Grand_Bargain_final_22_May_FINAL-2.pdf

World Bank. (2020). *World Bank Country and Lending Groups*. The World Bank. https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups

5.7. Chapter 5 - Appendix 1

Scoping Review Protocol Developed A Priori

This study uses the scoping review method established by Arksey and O'Malley (2005) and further refined by Levac et al (2010) and follows the framework maintained by the Joanna Briggs Institute (Peters et al., 2015, 2017). The protocol was revised based on feedback received from the research team and incorporated the results from a pilot conducted for this study November-December 2019. It follows the Preferred Reporting Items for Systematic Reviews and Meta-analysis for Protocols – Extension for Scoping Reviews (PRISMA-ScR) (Tricco et al., 2018) for complete and transparent publishing of scoping reviews. This protocol has been adapted from Tricco et al.(2016).

Review title and timescale

1 **Review title:**
Ethical Implications Related to Processing Personal Data in Humanitarian Crises: A Scoping Review

2 **Anticipated or actual start date:**
12/2/2019

3 **Anticipated completion date:**
8/1/2020

4 **Stage of review at time of this submission:**

Review stage	Started	Completed
Preliminary searches	☐	X
Piloting of the study selection process	☐	X
Formal screening of search results against eligibility criteria	X	☐
Data extraction	☐	☐
Risk of bias (quality) assessment	☐	N/A
Data analysis	☐	☐

Review team details

5 **Named contact**
Tino Kreutzer

6 **Named contact email**
kreutzer@yorku.ca

7 **Named contact address**
Suite 2150, Dahdaleh Building, York University
88 The Pond Road, Toronto, Ontario, Canada, M3J 1P3

8 **Named contact phone number**
+1 416-736-2100 x34447

9 **Organizational affiliation of the review**
Dahdaleh Institute for Global Health Research, York University

10 **Review team members and their organizational affiliations**
Mr. Tino Kreutzer, York University (PhD candidate)
Dr. James Orbinski, York University (dissertation supervisor)
Dr. Lora Appel, York University (dissertation committee member)
Dr. Aijun An, York University (dissertation committee member)
Dr. Patrick Vinck, Harvard University (dissertation committee member)

11 **Funding sources/sponsors**
N/A

12 **Conflicts of interest**
Authors have no known conflicts of interest to declare.

Review methods

13 **Review question(s):**
The objective of this scoping review is to map existing research on ethical implications stemming from the processing of data to inform humanitarian assistance. The specific research questions are:

Which **ethical implications** related to processing data from people affected by humanitarian crises in order to inform humanitarian assistance have been raised?
What was the empirical basis used to establish these implications?
Which technologies did these implications relate to?
What types of humanitarian crises did the study relate to?

14 **Literature Search:**
Comprehensive literature searches of electronic databases were conducted on March 31, 2020 using Ovid, Ebsco, Web of Science, and Proquest to search 20 databases for relevant studies. Only studies published in English, French, or Spanish were included. Search was limited to publications from 2010 onwards.

15 **URL to search strategy:**
See https://bit.ly/3fIqJgO for full details and search syntax.

16 **Condition or domain being studied:**

Ethical implications stemming from the processing of data relating to people affected by a humanitarian crisis with the explicit goal or potential of informing humanitarian assistance.

17 **Population/Problem:**

People affected by a **humanitarian crisis**.[1] This includes armed conflicts, natural disasters, and large public health emergencies—as well as refugees and migrants fleeing from such a crisis, regardless of their current location. We will also include studies that concern humanitarian assistance (including related fields such as disaster response or emergency management) that are global in scope. The following are included for the purpose of this study:

- Natural disasters in low or lower middle-income countries (defined as countries that ranked low income or low middle income at least once by the World Bank between 2011 and 2020) (World Bank, 2020)
- The Ebola outbreaks in West Africa (2014-2016) and DR Congo (2018-2020)
- Any events taking place in countries registered in the same year as recipients by UNOCHA's Financial Tracking Service (https://fts.unocha.org)

Excluded are:

- Studies related solely on natural disasters high or high middle-income countries

[1] Humanitarian crises are defined here as "an event or series of events representing a critical threat to the health, safety, security or wellbeing of a community, usually over a wide area" (WHO, 2007).

18 **Intervention(s)/Exposure(s):**
Studies that **investigate the processing of data**[2] **relating to people affected by a humanitarian crisis with the explicit goal or potential of informing humanitarian assistance**[3].

Excluded from this definition are:

- Studies that discuss data processing purely for research purposes
- Any technologies that do not collect or process data related to affected people, e.g., robotics for clearing debris or land mines or any other robotics
- Military aircraft (unless used to collect such data to support humanitarian assistance)
- Algorithmic models for predicting the occurrence or impacts of natural hazards
- Tools used for planning humanitarian logistics as long as they do not involve personal data from affected people (e.g., relief / distribution networks, supply chain management, and resource scheduling).

19 **Comparator(s)/Control(s):**
Studies without a comparator are eligible for inclusion. Any comparator is relevant for inclusion, such as comparing different technologies with each other or versus no technology used.

20 **Types of study to be included:**
All study designs will be eligible for inclusion.

Excluded are:

---

[2] Data processing is defined here as "Any operation or set of operations which is performed on data or on sets of data, whether or not by automated means, such as collecting, registering, storing, adapting or altering, cleaning, filing, retrieving, using, disseminating, transferring and retaining or destroying" (OCHA, 2019).

[3] Humanitarian assistance refers to coordinated actions that save lives, alleviate suffering, and maintain human dignity during and after human-made crises and disasters caused by natural hazards (Pictet, 1979). Humanitarian assistance here is considered to include "protection", which "encompasses all activities aimed at obtaining full respect for the rights of the individual in accordance with the letter and the spirit of the relevant bodies of law" (IASC, 1999, p. 4).

- Non peer-reviewed studies
- Congressional documents and publications (e.g., committee hearing reports)
- Conference announcements or summaries (e.g., "abstract list", "poster abstracts", etc.)
- Book reviews

21 **Context:**
For feasibility reasons, we restrict the review to studies published since 1/1/2010. Studies in all countries or territories affected by a humanitarian crisis (or relevant host countries for displaced populations) will be included, as defined in point 17.

22 **Primary outcome(s):**
Studies that **investigate ethical implications**[4] **stemming from the processing of such data** (as defined in point 18).

Only studies that contain a **meaningful discussion** about this subject will be included. This will be assessed initially by whether ethical issues are explicitly mentioned in the abstract (including by looking for a broad range of potential synonyms, such as challenges, problems, risks, etc.). During the full text review this will be assessed qualitatively by two reviewers.

Empirical studies and those developing theory or frameworks, as well as reports, commentary or other types of articles will be included.

23 **Secondary outcome(s):**
Not applicable.

24 **Data extraction (selection and coding):**
Study selection and coding will be done using the DistillerSR software. Using the a priori eligibility criteria (points 17, 18, 20, and 22), standardized questionnaires have been developed. Testing and training exercises will precede each level of screening. Reviewer pairs will screen references and full texts independently.

The following selection process will be used:

<u>Title screening</u>
- Two reviewers will independently select studies solely on their title and bibliographic information

[4] Ethical implications are defined here as actions that may not be conforming to moral standards, particularly those set out by various humanitarian principles (Slim, 2015, pp. 47–145) because of the risks they presented

- Daily meetings will be held to compare working definitions and discuss rating discrepancies during the first 1,000 references being reviewed
- Each reference may be excluded based on the following screening questions:
    1. Is the study likely to be about a humanitarian crisis?
    2. [if (1) = yes or unsure]: Is the study likely to investigate the processing of data relating to people affected by a humanitarian crisis?
    3. Is the type of study and study language eligible?
- A study is excluded if any question was answered 'no'
- Any conflicts during this title screening will be included in the abstract review, meaning that if one reviewer included and one excluded the reference it is considered included

<u>Abstract screening</u>

- Two reviewers will independently select studies based on their title, abstract, and bibliographic details
- Daily meetings will be held during the review of the first 100 references to discuss rating discrepancies
- Each reference may be excluded based on the abovementioned title screening questions, plus the following:
    1. [if (2) = yes or unsure]: Is the study likely to have a meaningful discussion of ethical implications?
- A study is excluded if any question was answered 'no'
- Any conflicts during this abstract screening will be included in the full text review, meaning that if one reviewer included and one excluded the reference it is considered included

<u>Full text review</u>

- Two reviewers will independently select studies based on the full text
- Daily meetings will be held during the review of the first 20 references to discuss rating discrepancies
- Each reference may be excluded based on the abovementioned screening criteria (see Appendix for the exact screening questions)
- A study is excluded if any question was answered 'no'
- Rating discrepancies will be resolved by discussion or by using a third adjudicator

<u>Data extraction</u>

- Two reviewers will independently collect data from each study. In case a large number of studies is identified (>30), we will conduct data extraction with one reviewer and one verifier.
- Data collection form will be pilot-tested prior to full data extraction
- We will extract data on
    1. Study characteristics (e.g., country of corresponding author, journal discipline)
    2. Population characteristics (e.g., type of humanitarian crisis)
    3. intervention characteristics (e.g., purpose of data processing, technologies described)

4. Outcomes (e.g., specific ethical implications identified, whether implications are based on empirical data or theoretical, details of empirical data, whether artificial intelligence or related technologies were mentioned, specific implementation and/or compliance methods that may have been proposed)

25 **Risk of bias (quality) assessment:**
No quality appraisal will be conducted as this is a scoping review. This is consistent with the framework proposed by Arksey and O'Malley, as well as the Joanna Briggs Institute guidance on conducting Scoping Reviews.

26 **Strategy for data synthesis:**
Results will be summarized quantitatively (using frequencies) and qualitatively (using descriptive analytics) to map and to identify gaps in the existing literature.

27 **Analysis of subgroups or subsets:**
Not applicable

Review general information

28 **Type of review**
Scoping review
29 **Language**
English
30 **Country**
Canada
31 **Other registration details**
Not applicable
32 **Reference and/or URL for published protocol**
Not applicable
33 **Dissemination plans:**
Do you intend to publish the review on completion?
Yes x No ☐
34 **Keywords**

35 **Details of any existing review of the same topic by the same authors.**
Not applicable
36 **Current review status**
Ongoing (study selection phase)

37 **Any additional information:**

Not applicable

38 **Details of final report/publication(s):**

Not applicable (review still in progress)

References

Arksey, H., & O'Malley, L. (2005). Scoping studies: Towards a methodological framework. *International Journal of Social Research Methodology*, *8*(1), 19–32. https://doi.org/10.1080/1364557032000119616

IASC. (1999). *Protection of Internally Displaced Persons*. https://interagencystandingcommittee.org/system/files/legacy_files/protection_of_internally_displaced_persons_inter_agency_standing_committee_policy_paper_0.pdf

Levac, D., Colquhoun, H., & O'Brien, K. K. (2010). Scoping studies: Advancing the methodology. *Implementation Science*, *5*(1). https://doi.org/10.1186/1748-5908-5-69

OCHA. (2019). *Data responsibility guidelines*. UN Office for the Coordination of Humanitarian Affairs. https://centre.humdata.org/wp-content/uploads/2019/03/OCHA-DR-Guidelines-working-draft-032019.pdf

Peters, M. D. J., Godfrey, C. M., Khalil, H., McInerney, P., Parker, D., & Soares, C. B. (2015). Guidance for conducting systematic scoping reviews. *International Journal of Evidence-Based Healthcare*, *13*(3), 141–146. https://doi.org/10.1097/XEB.0000000000000050

Peters, M. D. J., Godfrey, C., Mcinerney, P., Baldini Soares, C., Khalil, H., & Parker, D. (2017). Chapter 11: Scoping reviews. In E. Aromataris & Z. Munn (Eds.), *Joana Briggs Institute Reviewer's Manual* (Issue September, pp. 6–24). https://reviewersmanual.joannabriggs.org/

Pictet, J. (1979). *The fundamental principles of the Red Cross*.

http://www.ifrc.org/PageFiles/40669/Pictet Commentary.pdf

Slim, H. (2015). *Humanitarian Ethics: A Guide to the Morality of Aid in War and Disaster*. Oxford University Press.

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D. J., Horsley, T., Weeks, L., Hempel, S., Akl, E. A., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., Wilson, M. G., Garritty, C., … Straus, S. E. (2018). PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. *Annals of Internal Medicine*, *169*(7), 467. https://doi.org/10.7326/M18-0850

Tricco, A. C., Zarin, W., Rios, P., Pham, B., Straus, S. E., & Langlois, E. V. (2016). Barriers, facilitators, strategies and outcomes to engaging policymakers, healthcare managers and policy analysts in knowledge synthesis: a scoping review protocol. *BMJ Open*, *6*(12), e013929. https://doi.org/10.1136/bmjopen-2016-013929

WHO. (2007). *Risk reduction and emergency preparedness*. World Health Organization. https://apps.who.int/iris/bitstream/handle/10665/43736/9789241595896_eng.pdf

World Bank. (2020). *World Bank Country and Lending Groups*. The World Bank. https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups

5.8. Chapter 5 - Appendix 2

Search Strategy and Keywords Used for Each Database

Concept		OVID	Ebsco	Scopus	WOS	Proquest
			((	(TITLE-ABS-KEY (	(TS=(	(noft(
Humanitarian assistance	1	humanitarian*.tw.	TI humanitarian* or AB humanitarian* or SU humanitarian*	humanitarian*	humanitarian*	humanitarian*
	2	relief work.tw.	or TI "relief work" or AB "relief work" or SU "relief work"	or "relief work"	or "relief work"	or "relief work"
	3	aid work.tw.	or TI "aid work" or AB "aid work" or SU "aid work"	or "aid work"	or "aid work"	or "aid work"
	4	(disaster? adj (relief or response? or assistance)).tw.	or TI (disaster# N0 (relief or response$ or assistance)) or AB (disaster# N0 (relief or response$ or assistance)) or SU (disaster# N0 (relief or response$ or assistance))	or (disaster# N/0 (relief or response or assistance))	or (disaster$ NEAR/0 (relief or response or assistance))	or (disaster? NEAR/0 (relief or response$ or assistance))
	5	emergency relief.tw.	or TI "emergency relief" or AB "emergency relief"	or "emergency relief"	or "emergency relief"	or "emergency relief"
	6	((conflict? or war?) adj10 (human rights or public health)).tw.	or TI ((conflict# or war#) N10 ("human rights" or "public health")) or AB ((conflict# or war#) N10 ("human rights" or "public health"))	or ((conflict# or war#) N/10 ("human rights" or "public health"))	or ((conflict$ or war$) NEAR/10 ("human rights" or "public health"))	or ((conflict? or war?) NEAR/10 ("human rights" or "public health"))
	7	(ebola adj6 (west africa or sierra leone or liberia or guinea or 2014 or 2013)).tw.	or TI (ebola N6 ("west africa" or "sierra leone" or liberia or guinea or 2014 or 2013)) or AB (ebola N6 ("west africa" or "sierra leone" or liberia or guinea or 2014 or 2013))	or (ebola N/6 ("west africa" or "sierra leone" or liberia or guinea or 2014 or 2013))	or (ebola NEAR/6 ("west africa" or "sierra leone" or liberia or guinea or 2014 or 2013))	or (ebola NEAR/6 ("west africa" or "sierra leone" or liberia or guinea or 2014 or 2013))
	8	acute malnutrition.tw.	or TI "acute malnutrition" or AB "acute malnutrition"	or "acute malnutrition"	or "acute malnutrition"	or "acute malnutrition"
	9	(refugee* adj2 (camp* or assistance or population?)).tw.	or TI (refugee* N2 (camp* or assistance or population#)) or AB (refugee* N2 (camp* or assistance or population#))	or (refugee* N/2 (camp* or assistance or population#))	or (refugee* NEAR/2 (camp* or assistance or population$))	or (refugee* NEAR/2 (camp* or assistance or population?))
	10	(displace* adj2 (forced or forcibly or population? or human? or internal*)).tw.	or TI (displace* N2 (forced or forcibly or population# or human# or internal*)) or AB (displace* N2 (forced or forcibly or population# or human# or internal*))	or (displace* N/2 (forced or forcibly or population# or human# or internal*))	or (displace* NEAR/2 (forced or forcibly or population$ or human$ or internal*))	or (displace* NEAR/2 (forced or forcibly or population? or human? or internal*))
	11	(((population? or person* or communit*) adj3 affected) adj1 (conflict? or violence)).tw.	or TI (((population# or person* or communit*) N3 affected) N1 (conflict# or violence)) or AB (((population# or person* or communit*) N3 affected) N1 (conflict# or violence))	or (((population# or person* or communit*) N/3 affected) N/1 (conflict# or violence))	or (((population$ or person* or communit*) NEAR/3 affected) NEAR/1 (conflict$ or violence))	or (((population? or person* or communit*) NEAR/3 affected) NEAR/1 (conflict? or violence))
	12	or/ 1-11	)	)	)	)
			or (	or (	or (TS=	or ((
	13	(cris?s or emergenc* or disaster?).tw.	(TI (cris#s or emergenc* or disaster#) or AB (cris#s or emergenc* or disaster#)	TITLE-ABS-KEY(cris#s or emergenc* or disaster#)	(cris$s or emergenc* or disaster$)	noft(cris?s or emergenc* or disaster?)
			) and	and	and	and
	14	humanitarian*.af.	TX humanitarian*	ALL(humanitarian*)	ALL=humanitarian*	ft(humanitarian*
	15	13 and 14	)) and	)) and	)) and	)))) and
	16	12 or 15				
ICT for data collection	17	ict.tw.	(TI ict or AB ict	TITLE-ABS-KEY(ict	TS=(ict	noft(ict
	18	technolog*.tw.	or TI technolog* or AB technolog*	or technolog*	or technolog*	or technolog*
	19	((data or information) adj2 (system* or manage* or collection or analys?s or process*)).tw.	or TI ((data or information) N2 (system* or manage* or collection or analys#s or process*)) or AB ((data or information) N2 (system* or manage* or collection or analys#s or process*))	or ((data or information) N/2 (system* or manage* or collection or analys#s or process*))	or ((data or information) NEAR/2 (system* or manage* or collection or analys$s or process*))	or ((data or information) NEAR/2 (system* or manage* or collection or analys?s or process*))
	20	(blockchain or distributed ledger).tw.	or TI (blockchain or "distributed ledger") or AB (blockchain or "distributed ledger")	or (blockchain or "distributed ledger")	or (blockchain or "distributed ledger")	or (blockchain or "distributed ledger")
	21	(ai or artificial intelligence or machine learning or algorithm*).tw.	or TI (ai or "artificial intelligence" or "machine learning" or algorithm*) or AB (ai or "artificial intelligence" or "machine learning" or algorithm*)	or (ai or "artificial intelligence" or "machine learning" or algorithm*)	or (ai or "artificial intelligence" or "machine learning" or algorithm*)	or (ai or "artificial intelligence" or "machine learning" or algorithm*)
	22	biometric*.tw.	or TI biometric* or AB biometric*	or biometric*	or biometric*	or biometric*
	23	smartphone app*.tw.	or TI "smartphone app*" or AB "smartphone app*"	or "smartphone app*"	or "smartphone app*"	or "smartphone app*"
	24	remote sensing.tw.	or TI "remote sensing" or AB "remote sensing"	or "remote sensing"	or "remote sensing"	or "remote sensing"
	25	analytics.tw.	or TI analytics or AB analytics	or analytics	or analytics	or analytics
	26	digital*.tw.	or TI digital* or AB digital*	or digital*	or digital*	or digital*
	27	experimentation.tw.	or TI experimentation or AB experimentation	or experimentation	or experimentation	or experimentation
	28	automat*.tw.	or TI automat* or AB automat*	or automat*	or automat*	or automat*
	29	innovation?.tw.	or TI innovation# or AB innovation#	or innovation#	or innovation$	or innovation?
	30	remote management.tw.	or TI "remote management" or AB "remote management"	or "remote management"	or "remote management"	or "remote management"
	31	cyber.tw.	or TI cyber or AB cyber	or cyber	or cyber	or cyber
	32	big data.tw.	or TI "big data" or AB "big data"	or "big data"	or "big data"	or "big data"
	33	(sms or text messag* or interactive voice recognition or online survey*).tw.	or TI (sms or "text messag*" or "interactive voice recognition" or "online survey*") or AB (sms or "text messag*" or "interactive voice recognition" or "online survey*")	or (sms or "text messag*" or "interactive voice recognition" or "online survey*")	or (sms or "text messag*" or "interactive voice recognition" or "online survey*")	or (sms or "text messag*" or "interactive voice recognition" or "online survey*")
	34	(kobotoolbox or kobo or odk or open data kit).tw.	or TI (kobotoolbox or kobo or odk or "open data kit") or AB (kobotoolbox or kobo or odk or "open data kit")	or (kobotoolbox or kobo or odk or "open data kit")	or (kobotoolbox or kobo or odk or "open data kit")	or (kobotoolbox or kobo or odk or "open data kit")
	35	crowdsourc*.tw.	or TI crowdsourc* or AB crowdsourc*		or crowdsourc*	or crowdsourc*
	36	social media.tw.	or TI "social media" or AB "social media"		or "social media"	or "social media"
	37	crisis adj (informatics or data or map*).tw.	or TI "crisis N0 (informatics or data or map*)" or AB "crisis N0 (informatics or data or map*)"		or "crisis NEAR/0 (informatics or data or map*)"	or "crisis NEAR/0 (informatics or data or map*)"
	38	digiti?ation.tw.	or TI digiti#ation or AB digiti#ation		or digiti$ation	or digiti?ation
	39	datafication.tw.	or TI datafication or AB datafication		or datafication	or datafication
	40	or/ 17-39	)	)	)	)
			and	and	and	and
Ethical concerns	41	concern?.tw.	(TI concern# or AB concern#	TITLE-ABS-KEY(concern#	TS=(concern$	noft(concern?
	42	risk?.tw.	or TI risk# or AB risk#	or risk#	or risk$	or risk?
	43	challenge?.tw.	or TI challenge# or AB challenge#	or challenge#	or challenge$	or challenge?
	44	harm?.tw.	or TI harm# or AB harm#	or harm#	or harm$	or harm?
	45	privacy.tw.	or TI privacy or AB privacy	or privacy	or privacy	or privacy
	46	protection?.tw.	or TI protection# or AB protection#	or protection#	or protection$	or protection?
	47	humanitarian adj (principle? or standard? or guideline?).tw.	or TI humanitarian N0 (principle# or standard# or guideline#) or AB humanitarian N0 (principle# or standard# or guideline#)	or humanitarian N/0 (principle# or standard# or guideline#)	or humanitarian NEAR/0 (principle$ or standard$ or guideline$)	or humanitarian NEAR/0 (principle? or standard? or guideline?)
	48	problem?.tw.	or TI problem# or AB problem#	or problem#	or problem$	or problem?
	49	bias?.tw.	or TI bias# or AB bias#	or bias#	or bias$	or bias?
	50	ethic*.tw.	or TI ethic* or AB ethic*	or ethic*	or ethic*	or ethic*
	51	consequence?.tw.	or TI consequence# or AB consequence#		or consequence$	or consequence?
	52	critique?.tw.	or TI critique# or AB critique#		or critique$	or critique?
	53	insecurity.tw.	or TI insecurity or AB insecurity		or insecurity	or insecurity
	54	implications.tw.	or TI implications or AB implications		or implications	or implications
	55	peril?.tw.	or TI peril# or AB peril#		or peril$	or peril?
	56	impact?.tw.	or TI impact# or AB impact#		or impact$	or impact?
	57	or/ 41-56	)	)	)	) and pd(2010-2019)
			)			
		16 and 40 and 57				

5.9. Chapter 5 - Appendix 3

Results Table for Studies Discussing AI (n = 25)

	Results for Studies Discussing AI (n=25)	Count
Year of publication	2010	
	2012	
	2013	
	2014	3
	2015	3
	2016	4
	2017	2
	2018	5
	2019	8
Type of humanitarian crisis	Natural disaster	9
	Armed conflict	7
	Refugees or migrants who fled a humanitarian crisis	6
	Large public health emergency	7
	Not specified	
Purposes of data processing	Assessment (of needs, damage, etc.)	9
	Registration / case management	6
	Medical care or public health	7
	Forecasting / modeling / early warning	6
	Delivery of assistance	5
	Accountability (complaints, feedback collection, etc.)	3
	Cash transfer	2
	Human rights violations	2
	Logistics	1
	Search and rescue	4
	Other	0
	Not specified	0
Specific technologies described	Social media	15
	Crowdsourcing	10
	Mapping / GIS	13
	Big data	15
	SMS or private messaging software	8
	AI / algorithms / machine learning	25

	Results for Studies Discussing AI (n=25)	Count
	Satellite imagery	7
	Biometrics	7
	Information systems	5
	Unmanned aerial vehicle (UAV)	6
	Medical data	6
	Cash distribution	3
	Call data records	4
	Data storage	2
	Blockchain / distributed ledger technology	2
	Computer-assisted personal interviewing	0
	Not specified	7
	computer-assisted telephone interviewing (CATI)	0
Ethical Issues		
Autonomy	Lack of consent: Data is collected without informed consent	16
	Data agency: People do not have the right to control, access, or delete their data	9
	Lack of respect: People/communities are not treated with respect	11
	Autonomy: Unwillingness to share data does not lead to disadvantages (e.g., exclusion from assistance or protection	10
	Participation: People/communities are not involved in decisions to use of new/experimental technologies for collecting data	5
	Undisclosed use: Data may be used beyond purposes for which they were collected	1
	Lack of group agency: Processed information is not available to affected communities	0
Beneficence	Unreliability: Processed data is inaccurate and does not sufficiently reflect reality to inform assistance	17
	Dependence: Data is processed with the assistance of a political, economic, or military entity	14
	Lack of action: Processed data is not utilized to inform assistance to the affected person/community	11
	Non-neutrality: Data is processed in a way that benefits or appears to benefit one side of the conflict over the other	8
	Ineffective or inefficient: Not producing expected result, unmet expectations	1
Non-maleficence	Privacy: Personal/sensitive data is shared with third parties	23
	Harm: People may suffer physical or psychological harm as a result of data processing	22
	Data security: Personal/sensitive data is not protected against malicious actors	8

	Results for Studies Discussing AI (n=25)	Count
	Power imbalance: Data processing reinforces or worsens a lack of power of affected people	8
	Excess: More data was collected than necessary	5
	Redress/rectification: People do not have the ability to correct wrong information about them or receive compensation	3
Justice	Bias: Data is processed in a way that may result in aid being distributed disproportionate with regard to people's actual needs	18
	Lack of accountability: Endangering (or not protecting) rights; absolving responsibility	6
	Unequal access to technology / exclusion from data collection	3
	Unfair distribution of risks and benefits	1

5.10. Chapter 5 - Appendix 4

List of Studies Included in Scoping Review Results

Beduschi, A. (2019). Digital identity: Contemporary challenges for data protection, privacy and non-discrimination rights. *Big Data and Society*, *6*(2), 1–6. https://doi.org/10.1177/2053951719855091

Belliveau, J. (2016). Humanitarian access and technology: Opportunities and applications. *Procedia Engineering*, *159*(June), 300–306. https://doi.org/10.1016/j.proeng.2016.08.182

Bempong, N. E., De Castañeda, R. R., Schütte, S., Bolon, I., Keiser, O., Escher, G., & Flahault, A. (2019). Precision Global Health - The case of Ebola: A scoping review. *Journal of Global Health*, *9*(1). https://doi.org/10.7189/jogh.09.010404

Bennett, C. (2014). Who knows who we are? Questioning DNA analysis in disaster victim identification. *New Genetics and Society*, *33*(3), 239–256. https://doi.org/10.1080/14636778.2014.946003

Bernard, R., Bowsher, G., Milner, C., Boyle, P., Patel, P., & Sullivan, R. (2018). Intelligence and global health: assessing the role of open source and social media intelligence analysis in infectious disease outbreaks. *Journal of Public Health (Germany)*, *26*(5), 509–514. https://doi.org/10.1007/s10389-018-0899-3

Bittner, C., Glasze, G., & Turk, C. (2013). Tracing contingencies: analyzing the political in assemblages of web 2.0 cartographies. *GeoJournal*, *78*(6), 935. https://doi.org/http://dx.doi.org/10.1007/s10708-013-9488-8

Bock, J. G. (2015). Firmer Footing for a Policy of Early Intervention: Conflict Early Warning and Early Response Comes of Age. *Journal of Information Technology and Politics*, *12*(1),

103–111. https://doi.org/10.1080/19331681.2014.982265

Bock, J. G. (2016). Technology and Vulnerability in Early Warning: Ethical Use of IT in Dangerous Places. *Information Technology for Development*, *22*(4), 696–709. https://doi.org/10.1080/02681102.2014.903894

Bravo, R. Z. B., Leiras, A., & Cyrino Oliveira, F. L. (2019). The Use of UAVs in Humanitarian Relief: An Application of POMDP-Based Methodology for Finding Victims. *Production & Operations Management*, *28*(2), 421–440. http://10.0.4.87/poms.12930

Brown, C. (2018). The use of ICTs in conflict and peacebuilding: A feminist analysis. *Australian Feminist Law Journal*, *44*(1), 137–153. https://doi.org/10.1080/13200968.2018.1465332

Burger, A., Oz, T., Kennedy, W. G., & Crooks, A. T. (2019). Computational social science of disasters: Opportunities and challenges. *Future Internet*, *11*(5), 1–31. https://doi.org/10.3390/fi11050103

Burns, R. (2014). Moments of closure in the knowledge politics of digital humanitarianism. *Geoforum*, *53*, 51–62. http://10.0.3.248/j.geoforum.2014.02.002

Burns, R. (2019). New Frontiers of Philanthro-capitalism: Digital Technologies and Humanitarianism. *Antipode*, *51*(4), 1101–1122. https://doi.org/10.1111/anti.12534

Carenzo, L., Barra, F. L., Ingrassia, P. L., Colombo, D., Costa, A., & Corte, F. Della. (2015). Disaster medicine through Google Glass. *European Journal of Emergency Medicine*, *22*(3), 222–225. https://doi.org/10.1097/MEJ.0000000000000229

Chan, J., Bateman, L., & Olafsson, G. (2016). A people & purpose approach to humanitarian data information security and privacy. *Procedia Engineering*, *159*(June), 3–5. https://doi.org/10.1016/j.proeng.2016.08.056

Cinnamon, J., Jones, S. K., & Adger, W. N. (2016). Evidence and future potential of mobile

phone data for disease disaster management. *Geoforum*, *75*, 253–264. https://doi.org/10.1016/j.geoforum.2016.07.019

Clark, N., & Guiffault, F. (2018). Seeing through the clouds: Processes and challenges for sharing geospatial data for disaster management in Haiti. *International Journal of Disaster Risk Reduction*, *28*(February), 258–270. https://doi.org/10.1016/j.ijdrr.2018.02.019

Comes, T. (2016). Cognitive biases in humanitarian sensemaking and decision-making lessons from field research. *2016 IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support, CogSIMA 2016*, *March*, 56–62. https://doi.org/10.1109/COGSIMA.2016.7497786

Comes, T., Meesters, K., & Torjesen, S. (2019). Making sense of crises: the implications of information asymmetries for resilience and social justice in disaster-ridden communities. *Sustainable and Resilient Infrastructure*, *4*(3), 124–136. https://doi.org/10.1080/23789689.2017.1405653

Crawford, K., & Finn, M. (2015). The limits of crisis data: analytical and ethical challenges of using social and mobile data to understand disasters. *GeoJournal*, *80*(4), 491–502. https://doi.org/10.1007/s10708-014-9597-z

Curry, T., Croitoru, A., Crooks, A., & Stefanidis, A. (2019). Exodus 2.0: crowdsourcing geographical and social trails of mass migration. *Journal of Geographical Systems*, *21*(1), 161–187. http://10.0.3.239/s10109-018-0278-1

Duffield, M. (2016). The resilience of the ruins: towards a critique of digital humanitarianism. *Resilience*, *4*(3), 147–165. https://doi.org/10.1080/21693293.2016.1153772

Easton, C. (2017). Analysing the Role of Privacy Impact Assessments in Technological Development for Crisis Management. *Journal of Contingencies & Crisis Management*,

25(1), 7–14. http://10.0.4.87/1468-5973.12140

Fast, L. (2017). Diverging Data: Exploring the Epistemologies of Data Collection and Use among Those Working on and in Conflict. *International Peacekeeping*, *24*(5), 706–732. https://doi.org/10.1080/13533312.2017.1383562

Garattini, C., Raffle, J., Aisyah, D. N., Sartain, F., & Kozlakidis, Z. (2019). Big Data Analytics, Infectious Diseases and Associated Ethical Impacts. *Philosophy and Technology*, *32*(1), 69–85. https://doi.org/10.1007/s13347-017-0278-y

Halkort, M. (2016). Liquefying Social Capital On the Bio-politics of Digital Circulation in a Palestinian Refugee Camp. *Tecnoscienza-Italian Journal of Science & Technology Studies*, *7*(2), 61–79.

Halkort, M. (2019). Decolonizing data relations: On the moral economy of data sharing in Palestinian refugee camps. *Canadian Journal of Communication*, *44*(3), 317–329. https://doi.org/10.22230/cjc.2019v44n3a3457

Hassan, N. H., & Rahim, F. A. (2017). The rise of crowdsourcing using social media platforms: Security and privacy issues. *Pertanika Journal of Science and Technology*, *25*(S10), 79–88.

Haworth, B. T., Bruce, E., Whittaker, J., & Read, R. (2018). The good, the bad, and the uncertain: Contributions of volunteered geographic information to community disaster resilience. *Frontiers in Earth Science*, *6*(November), 1–15. https://doi.org/10.3389/feart.2018.00183

Hayes, B. (2017). Migration and data protection: Doing no harm in an age of mass displacement, mass surveillance and “big data.” *International Review of the Red Cross*, *99*(904), 179. https://doi.org/http://dx.doi.org/10.1017/S1816383117000637

Hayes, P., & Kelly, S. (2018). Distributed morality, privacy, and social media in natural disaster

response. *Technology in Society*, *54*(December 2017), 155–167. https://doi.org/10.1016/j.techsoc.2018.05.003

Hudson, L. W. (2013). SMS-based accountability to beneficiaries mechanisms in humanitarian aid and development. *ACM International Conference Proceeding Series*, *2 NOTES*, 167–170. https://doi.org/10.1145/2517899.2517942

Hunt, M., Pringle, J., Christen, M., Eckenwiler, L., Schwartz, L., & Davé, A. (2016). Ethics of emergent information and communication technology applications in humanitarian medical assistance old. *International Health*, *8*(4), 239–245. https://doi.org/10.1093/inthealth/ihw028

Jacobsen, K. L. (2010). Making design safe for citizens: a hidden history of humanitarian experimentation. *Citizenship Studies*, *14*(1), 89. https://doi.org/http://dx.doi.org/10.1080/13621020903466399

Jacobsen, K. L. (2015). Experimentation in humanitarian locations: UNHCR and biometric registration of Afghan refugees. *Security Dialogue*, *46*(2), 144–164. https://doi.org/10.1177/0967010614552545

Jacobsen, K. L. (2017). On Humanitarian Refugee Biometrics and New Forms of Intervention. *Journal of Intervention and Statebuilding*, *11*(4), 529–551. https://doi.org/10.1080/17502977.2017.1347856

Jacobsen, K. L., & Fast, L. (2019). Rethinking access: how humanitarian technology governance blurs control and care. *Disasters*, *43*(S2), S151–S168. https://doi.org/10.1111/disa.12333

Jacobsen, K. L., & Sandvik, K. B. (2018). UNHCR and the pursuit of international protection: accountability through technology? *Third World Quarterly*, *39*(8), 1508–1524. http://10.0.4.56/01436597.2018.1432346

Kroener, I., Barnard-Wills, D., & Muraszkiewicz, J. (2019). Agile ethics: an iterative and flexible approach to assessing ethical, legal and social issues in the agile development of crisis management information systems. *Ethics and Information Technology*, 1. https://doi.org/http://dx.doi.org/10.1007/s10676-019-09501-6

Kunz, N. (2019). An automated quantitative content analysis process for humanitarian logistics research. *Journal of Humanitarian Logistics and Supply Chain Management*, *9*(3), 475–491. https://doi.org/10.1108/JHLSCM-06-2018-0051

Lambert, A. (2016). Disaster data assemblages: five perspectives on social media and communities in response and recovery. *Proceedings of the Annual Hawaii International Conference on System Sciences*, *2016-March*, 2237–2245. https://doi.org/10.1109/HICSS.2016.280

Latif, S., Qayyum, A., Usama, M., Qadir, J., Zwitter, A., & Shahzad, M. (2019). Caveat Emptor: The Risks of Using Big Data for Human Development. *IEEE Technology & Society Magazine*, *38*(3), 82–90. http://10.0.4.85/MTS.2019.2930273

Le Blond, S., Cuevas, A., Troncoso-Pastoriza, J. R., Jovanovic, P., Ford, B., & Hubaux, J. P. (2018). On Enforcing the Digital Immunity of a Large Humanitarian Organization. *Proceedings - IEEE Symposium on Security and Privacy*, *2018-May*, 424–440. https://doi.org/10.1109/SP.2018.00019

Lev Aretz, Y. (2019). Data Philanthropy. *SSRN Electronic Journal*, *70*(6). https://doi.org/10.2139/ssrn.3320798

Longboan, L. (2018). “I don’t want trouble”: Freedom of expression and use of backchannels in the wake of Typhoon Haiyan. *Disaster Prevention and Management: An International Journal*, *27*(4), 380–392. https://doi.org/10.1108/DPM-01-2018-0021

Madianou, M. (2015). Digital Inequality and Second-Order Disasters: Social Media in the Typhoon Haiyan Recovery. *Social Media and Society*, *1*(2). https://doi.org/10.1177/2056305115603386

Madianou, M. (2019). The Biometric Assemblage: Surveillance, Experimentation, Profit, and the Measuring of Refugee Bodies. *Television and New Media*, *20*(6), 581–599. https://doi.org/10.1177/1527476419857682

Madianou, M., Ong, J. C., Longboan, L., & Cornelio, J. S. (2016). The Appearance of Accountability: Communication Technologies and Power Asymmetries in Humanitarian Aid and Disaster Recovery. *Journal of Communication*, *66*(6), 960–981. http://10.0.4.87/jcom.12258

Madon, S., & Schoemaker, E. (2019). Reimagining Refugee Identity Systems: A Sociological Approach. *Information and Communication Technologies for Development: Strengthening Southern-Driven Cooperation as a Catalyst for Ict4d, Pt I*.

Martin-Shields, C. (2013). The Technologists Dilemma: Ethical Challenges of Using Crowdsourcing Technology in Conflict and Disaster-Affected Regions. *Georgetown Journal of International Affairs*, *14*(2), 157–163. http://www.jstor.com/stable/43134422

Martin, R. S., & Painho, M. (2019). Geospatial preparedness: Empirical study of the joint effort to provide geospatial support to disaster response. *Transactions in GIS*, *23*(3), 481. https://doi.org/http://dx.doi.org/10.1111/tgis.12537

Meier, P., & Munro, R. (2010). The Unprecedented Role of SMS in Disaster Response: Learning from Haiti. *Part of a Special Issue: The Cyber Challenge: Threats and Opportunities in a Networked World*, *30*(2), 91–103. http://ezproxy.library.yorku.ca/login?url=http://search.ebscohost.com/login.aspx?direct=tru

e&db=ssa&AN=510895864&site=ehost-live

Mesmar, S., Talhouk, R., Akik, C., Olivier, P., Elhajj, I. H., Elbassuoni, S., Armoush, S., Kalot, J., Balaam, M., Germani, A., & Ghattas, H. (2016). The impact of digital technology on health of populations affected by humanitarian crises: Recent innovations and current gaps. *Journal of Public Health Policy*, *37*(2), S167–S200. https://doi.org/10.1057/s41271-016-0040-1

Molnar, P. (2019). New technologies in migration: human rights impacts. *Forced Migration Review*, *61*, 7. http://ezproxy.library.yorku.ca/login?url=https://search.proquest.com/docview/2263276132?accountid=15182

Mulder, F., Ferguson, J., Groenewegen, P., Boersma, K., & Wolbers, J. (2016). Questioning Big Data: Crowdsourcing crisis data towards an inclusive humanitarian response. *Big Data and Society*, *3*(2), 1–13. https://doi.org/10.1177/2053951716662054

Palmer, L. (2014). Ushahidi at the Google interface: critiquing the 'geospatial visualization of testimony'. *Continuum: Journal of Media & Cultural Studies*, *28*(3), 342–356. http://10.0.4.56/10304312.2014.893989

Paul, S., & Sosale, S. (2020). Witnessing a Disaster: Public Use of Digital Technologies in the 2015 South Indian Floods. *Digital Journalism*, *8*(1), 15–31. https://doi.org/10.1080/21670811.2019.1636693

Peckham, R., & Sinha, R. (2017). Satellites and the New War on Infection: Tracking Ebola in West Africa. *Geoforum*, *80*, 24. http://ezproxy.library.yorku.ca/login?url=https://search.proquest.com/docview/1931697398?accountid=15182

Perakslis, E. D. (2018). Using digital health to enable ethical health research in conflict and other humanitarian settings. *Conflict and Health*, *12*(1), 23. https://doi.org/10.1186/s13031-018-0163-z

Perera, S. (2017). To Boldly Know: Knowledge, Peacekeeping and Remote Data Gathering in Conflict-Affected States. *International Peacekeeping*, *24*(5), 803–822. https://doi.org/10.1080/13533312.2017.1383566

Pham, P. N., & Vinck, P. (2012). Technology fusion and their implications for conflict early warning systems, public health, and human rights. *Health and Human Rights*, *14*(2), 106–117. http://www.jstor.org/stable/healhumarigh.14.2.106

Phillips, J. (2015). Exploring the citizen-driven response to crisis in cyberspace, risk and the need for resilience. *2015 IEEE Canada International Humanitarian Technology Conference, IHTC 2015*. https://doi.org/10.1109/IHTC.2015.7238051

Phillips, J. (2018). Risk in a digital age: understanding risk in virtual networks through digital response networks (DRNs). *International Development Planning Review*, *40*(3), 239–272. https://doi.org/10.3828/idpr.2018.18

Poblet, M., & Kolieb, J. (2018). Responding to Human Rights Abuses in the Digital Era: New Tools, Old Challenges. *Stanford Journal of International Law*, *54*(2), 259–283. http://ezproxy.library.yorku.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=130412330&site=ehost-live

Pond, P. (2016). The space between us: Twitter and crisis communication. *International Journal of Disaster Resilience in the Built Environment*, *7*(1), 40–48. https://doi.org/10.1108/IJDRBE-08-2013-0030

Ray, A., & Bala, P. K. (2020). Social media for improved process management in organizations

during disasters. *Knowledge & Process Management*, *27*(1), 63–74. http://10.0.3.234/kpm.1623

Raymond, N. A. (2017). Beyond "Do No Harm" and Individual Consent: Reckoning with the Emerging Ethical Challenges of Civil Society's Use of Data. In B. Taylor, L and Floridi, L and VanDerSloot (Ed.), *Group Privacy: New Challenges of Data Technologies* (Vol. 126, pp. 67–82). Springer. https://doi.org/10.1007/978-3-319-46608-8_4

Read, R. (2016). Tensions in un Information Management: Security, Data and Human Rights Monitoring in Darfur, Sudan. *Journal of Human Rights Practice*, *8*(1), 101–115. https://doi.org/10.1093/jhuman/huv018

Read, R., Taithe, B., & Mac Ginty, R. (2016). Data hubris? Humanitarian information systems and the mirage of technology. *Third World Quarterly*, *37*(8), 1314–1331. https://doi.org/10.1080/01436597.2015.1136208

Resor, E. (2016). The Neo-Humanitarians: Assessing the Credibility of Organized Volunteer Crisis Mappers. *Policy and Internet*, *8*(1), 34. https://doi.org/http://dx.doi.org/10.1002/poi3.112

Roberts, T., & Marchais, G. (2018). ASSESSING THE ROLE OF SOCIAL MEDIA AND DIGITAL TECHNOLOGY IN VIOLENCE REPORTING. *Contemporary Readings in Law and Social Justice*, *10*(2), 9–42. https://doi.org/http://dx.doi.org/10.22381/CRLSJ10220181

Sandvik, K. B. (2016). The humanitarian cyberspace: shrinking space or an expanding frontier? *Third World Quarterly*, *37*(1), 17. https://doi.org/http://dx.doi.org/10.1080/01436597.2015.1043992

Sandvik, K. B., & Lohne, K. (2014). The Rise of the Humanitarian Drone: Giving Content to an Emerging Concept. *Millennium (03058298)*, *43*(1), 145–164.

http://10.0.4.153/0305829814529470

Sandvik, K. B., Gabrielsen Jumbert, M., Karlsrud, J., & Kaufmann, M. (2014). Humanitarian technology: A critical research agenda. *International Review of the Red Cross*, *96*(893), 219–242. https://doi.org/10.1017/S1816383114000344

Sandvik, K. B., Jacobsen, K. L., & McDonald, S. M. (2017). Do no harm: A taxonomy of the challenges of humanitarian experimentation. *International Review of the Red Cross*, *99*(904), 319–344. https://doi.org/10.1017/S181638311700042X

Scarnecchia, D. P., Raymond, N. A., Greenwood, F., Howarth, C., & Poole, D. E. (2017). A Rights-based Approach to Information in Humanitarian Assistance. *PLoS Currents*, 1–5. https://doi.org/10.1371/currents.dis.dd709e442c659e97e2583e0a9986b668

Schoemaker, E., Baslan, D., Pon, B., & Dell, N. (2020). Identity at the margins: data justice and refugee experiences with digital identity systems in Lebanon, Jordan, and Uganda. *Information Technology for Development*, 13–36. https://doi.org/10.1080/02681102.2020.1785826

Schopper, D., Dawson, A., Upshur, R., Ahmad, A., Jesani, A., Ravinetto, R., Segelid, M. J., Sheel, S., & Singh, J. (2015). Innovations in research ethics governance in humanitarian settings. *BMC Medical Ethics*, *16*(1), 1–12. https://doi.org/10.1186/s12910-015-0002-3

Shah, D. C., Anderson, C., Breimyer, P., Foster, S., Geyer, K., Griffith, J. D., Heier, A., Majumdar, A., Simek, O., Stanisha, N., & Waugh, F. R. (2015). Application of graph methods for leveraging open source data during disaster response. *Proceedings of the 5th IEEE Global Humanitarian Technology Conference, GHTC 2015*, 259–266. https://doi.org/10.1109/GHTC.2015.7343982

Shanley, L. A., Burns, R., Bastian, Z., & Robson, E. S. (2013). Tweeting up a storm: The

promise and perils of crisis mapping. *Photogrammetric Engineering and Remote Sensing*, *79*(10), 865–879. https://doi.org/10.2139/ssrn.2464599

Sheather, J., Jobanputra, K., Schopper, D., Pringle, J., Venis, S., Wong, S., & Vincent-Smith, R. (2016). A Médecins Sans Frontières Ethics Framework for Humanitarian Innovation. *PLoS Medicine*, *13*(9), 1–10. https://doi.org/10.1371/journal.pmed.1002111

Shemak, A. (2014). The cartographic dimensions of humanitarianism: Mapping refugee spaces in postearthquake Haiti. *Cultural Dynamics*, *26*(3), 251–275. https://doi.org/10.1177/0921374014527920

Sivanesan, S. (2019). Alex & I: narrative and network resistance. *Social Identities*, *25*(4), 559. https://doi.org/http://dx.doi.org/10.1080/13504630.2018.1514161

Tanzi, T. J., Chandra, M., Isnard, J., Camara, D., Sebastien, O., & Harivelo, F. (2016). Towards "Drone-Borne" Disaster Management: Future Application Scenarios. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, *III–8*, 181–189. https://doi.org/10.5194/isprsannals-iii-8-181-2016

Tanzi, T., Apvrille, L., Dugelay, J. L., & Roudier, Y. (2014). UAVs for humanitarian missions: Autonomy and reliability. *Proceedings of the 4th IEEE Global Humanitarian Technology Conference, GHTC 2014*, 271–278. https://doi.org/10.1109/GHTC.2014.6970292

Tatham, P., Ball, C., Wu, Y., & Diplas, P. (2017). Long-endurance remotely piloted aircraft systems (LE-RPAS) support for humanitarian logistic operations. *Journal of Humanitarian Logistics and Supply Chain Management*, *7*(1), 2–25. https://doi.org/http://dx.doi.org/10.1108/JHLSCM-05-2016-0018

Tatsidou, E., Tsiamis, C., Karamagioli, E., Boudouris, G., Pikoulis, A., Kakalou, E., & Pikoulis, E. (2019). Reflecting upon the humanitarian use of unmanned aerial vehicles (drones).

Swiss Medical Weekly, *149*(13–14), 1–6. https://doi.org/10.4414/smw.2019.20065

Taylor, L. (2016). The ethics of big data as a public good: which public? Whose good? *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *374*(2083), 20160126. https://doi.org/10.1098/rsta.2016.0126

Taylor, L., & Schroeder, R. (2015). Is bigger better? The emergence of big data as a tool for international development policy. *GeoJournal*, *80*(4), 503–518. https://doi.org/10.1007/s10708-014-9603-5

Thylin, T., & Duarte, M. F. N. (2019). Leveraging blockchain technology in humanitarian settings - opportunities and risks for women and girls. *Gender and Development*, *27*(2), 317. https://doi.org/http://dx.doi.org/10.1080/13552074.2019.1627778

van Wynsberghe, A., & Comes, T. (2020). Drones in humanitarian contexts, robot ethics, and the human–robot interaction. *Ethics and Information Technology*, *22*(1), 43–53. https://doi.org/10.1007/s10676-019-09514-1

Vannini, S., Gomez, R., & Newell, B. C. (2019). Privacy and Security Guidelines for Humanitarian Work with Undocumented Migrants. *ICTD '19: Proceedings of the Tenth International Conference on Information and Communication Technologies and Development*. https://doi.org/10.1145/3287098.3287120

Vannini, S., Gomez, R., & Newell, B. C. (2020). "Mind the five": Guidelines for data privacy and security in humanitarian work with undocumented migrants and other vulnerable populations. *Journal of the Association for Information Science and Technology*, *71*(8), 927–938. https://doi.org/10.1002/asi.24317

Vannini, S., Gomez, R., Lopez, D., Mora, S., Morrison, J. C., Tanner, J., Youkhana, L., Vergara, G., & del Mar Moreno Tafurt, M. (2020). Humanitarian organizations' information

practices: Procedures and privacy concerns for serving the undocumented. *Electronic Journal of Information Systems in Developing Countries*, *86*(1), 1–8. https://doi.org/10.1002/isd2.12109

Venakata Swamy, M., Agarwal, N., & Ramaswamy, S. (2014). C-CBPM: collective context based privacy model. *Journal of Ambient Intelligence and Humanized Computing*, *5*(6), 881–895. https://doi.org/10.1007/s12652-014-0241-z

Voigt, S., Schoepfer, E., Fourie, C., & Mager, A. (2014). Towards semi-automated satellite mapping for humanitarian situational awareness. *Proceedings of the 4th IEEE Global Humanitarian Technology Conference, GHTC 2014*, 412–416. https://doi.org/10.1109/GHTC.2014.6970315

Watson, H., & Rodrigues, R. (2018). Bringing Privacy into the Fold: Considerations for the Use of Social Media in Crisis Management. *Journal of Contingencies & Crisis Management*, *26*(1), 89–98. http://10.0.4.87/1468-5973.12150

Wozniak, S., Rossberg, M., & Schaefer, G. (2013). Towards Trustworthy Mobile Social Networking Services for Disaster Response. *2013 IEEE International Conference on Pervasive Computing and Communications Workshops (Percom Workshops)*. http://arxiv.org/abs/1212.4989

Chapter 6 Systematic Design and Evaluation of New Humanitarian Needs Assessments Tools for Collecting Qualitative Data Using Natural Language Processing

Abstract

Interview data from humanitarian needs assessments (HNA) are commonly collected by entering responses into computer-assisted personal interviewing (CAPI) tools. However, the proper transcription and translation of qualitative interview responses are hard to conduct rapidly and at scale during a humanitarian crisis. This study reports on the design and empirical evaluation of new features for KoboToolbox, a CAPI tool supported and used by humanitarian organizations, to systematically transcribe, translate, and analyze answers to open-ended questions from HNA interviews. First, 23 interviews were conducted with international key informants (KIs) to explore accepted approaches to conducting HNA, as well as constraints related to different interview methods. Second, a scoping review that identified 100 studies, was conducted to map the range of ethical issues that had been raised in the peer-reviewed academic literature related to processing data of people affected by humanitarian crises. Third, using a user-centered design approach, new software features were designed and implemented, and usability testing was conducted based on observations and structured feedback from 14 test users. Interviews with 23 KIs and a review of 100 studies identified through the scoping review led to 25 themes and seven ethical issues, respectively, which were then used to inform the iterative design process. User testing showed strong support for the applicability and usability of new features and was subsequently used to identify and address several usability issues. It was concluded that the adoption of a multi-disciplinary user-centered design approach was crucial for

creating a new set of tools for scaling up the use of qualitative data in humanitarian needs assessments, while accommodating both the ethical realities and the operational constraints experienced in humanitarian assistance.

Systematic Design and Evaluation of New Humanitarian Needs Assessments Tools for Collecting Qualitative Data Using Natural Language Processing

Organizations involved in providing humanitarian assistance and responding to global health emergencies work under extreme conditions and with limited funding to provide life-saving support to people with the greatest need. In 2021, the cost of humanitarian assistance amounted to US $ 41 billion in order to assist as many as possible of the 274 million people requiring external aid (OCHA, 2021a). With increasingly inadequate financial resources to address growing humanitarian needs, Humanitarian needs assessments (HNAs) can provide critical evidence of population needs, so that the most appropriate interventions can be chosen; create a baseline for measuring their impacts, and can help prioritize use of limited funding (Banatvala, 2000; Cosgrave, 2009; Gerdin et al., 2014; Redmond, 2005; Spiegel et al., 2001). Both quantitative and qualitative approaches to collecting primary data to inform HNA are important, but significant technical, organizational, and systemic challenges make collecting qualitative data very difficult in practice, as described in Chapter 4 (Kreutzer, 2021b).

The aim of this study was to investigate how new software might be designed to enable specialists involved in managing HNA data to systematically transcribe, translate, and analyze answers to open-ended questions (OEQ) from HNA interviews. This section begins by providing an overview of the operational context for collecting HNA data. Second, it describes the needs expressed by HNA professionals to better manage qualitative data at scale, which led to the design proposal for integrating natural language processing (NLP) into the KoboToolbox software. Third, it briefly outlines three challenges that have been identified in relation to implementing such software features, namely limitations of NLP related to languages spoken in

humanitarian crises, ethical issues stemming from designing new data processing tools for humanitarian assistance, and usability issues that can limit implementation rates of new software.

Humanitarian assistance here refers to coordinated actions that save lives, alleviate human suffering, and maintain human dignity during and after human-made crises and disasters caused by natural hazards (IASC, 1999; Pictet, 1979). For the purposes of this study, humanitarian crises are defined as a "series of events representing a critical threat to the health, safety, security or wellbeing of a community, usually over a wide area" (WHO, 2007, p. 7). HNA are understood here to be the "set of activities necessary to understand a given situation, [which] entails the collection, updating and analysis of data pertaining to the population of concern (needs, capacities, resources, etc.), as well as the state of infrastructure and general socio-economic conditions in a given location/area" (UNHCR, 2006, p. 4). Data processing is defined here as any "operation or set of operations which is performed on data or on sets of data, whether or not by automated means, such as collecting, registering, storing, adapting or altering, cleaning, filing, retrieving, using, disseminating, transferring and retaining or destroying" (OCHA, 2019). NLP is understood here as "the use of computational methods to analyze and process spoken or written statements in a language commonly used by humans" (Assal et al., 2011, p. 2).

Interview data from HNA are commonly collected by entering responses into handheld computer-assisted personal interviewing (CAPI) tools. KoboToolbox (Kobo, 2022), a CAPI tool created in collaboration with the UN Office for the Coordination of Humanitarian Affairs (OCHA) and the UN High Commissioner for Refugees (UNHCR), is used by most humanitarian organizations for HNA and similar purposes (Building Markets & Orange Door Research, 2018).

However, such CAPI tools are primarily effective with quantitative data collection using structured interviews that use standardized questionnaires in which primarily closed-ended questions are asked in a pre-determined order. Yet, mixed-methods approaches that combine qualitative and quantitative data collection methods are often considered optimal for informing relevant humanitarian assistance during in-depth assessments (OCHA, 2021b; UNHCR, 2017). In practice, the collection and reliable analysis of qualitative data are considered too challenging, costly, and slow for inclusion in HNA—particularly during the acute emergency phase of a humanitarian crisis.

The urgent need to address this issue was first discussed by several leading humanitarian organizations at the International Health Emergency Data Science Workshop at York University, 4-5 December 2018. A working group convened by the Dahdahleh Institute of Global Health Research at with representatives from ACAPS, Elrha, Harvard Humanitarian Initiative (HHI) , the International Rescue Committee (IRC), NetHope, OCHA, Pivotal, Purple Compass, World Food Programme (WFP), and the Dahdaleh Institute for Global Health Research at York University (Emergency Data Science Workshop, 2018). Several members of this working group collaborated to describe the potential role that natural language processing (NLP), a type of artificial intelligence (AI), can play in increasing the speed and scope for processing qualitative interview data, as shown in Chapter 3 (Kreutzer et al., 2020). The working group concluded that focusing on NLP innovations that can save valuable time and costs while making richer data available to humanitarian organizations—especially by utilizing tools such as AI—carries enormous potential (Emergency Data Science Workshop, 2018). However, three distinct challenges were noted: 1) NLP is not available for many languages spoken during humanitarian

crises; 2) there are many ethical issues that arise when designing new data processing tools for humanitarian assistance, and 3) usability issues often limit implementation rates, particularly in non-commercial sectors. Each of these is discussed below.

First, in recent years, NLP has surpassed several important milestones (Galanis et al., 2021), particularly due to breakthroughs in the areas of deep learning and neural networks (Otter et al., 2021). The rapid advances in these fields (see, for example, Bansal et al., 2018; Weiss et al., 2017) have led to a boom in commercial applications for automated speech recognition (ASR) and machine translation (MT), as well as different types of content analysis (e.g., from Google, IBM Watson, Microsoft Azure, Amazon Web Services). At the same time, open-source software and training models are being created to replicate or surpass the performance of these systems. Open-source examples include Transformer (Vaswani et al., 2017), DeepSpeech (Hannun et al., 2014), and OpenNMT (Klein et al., 2017). But despite this progress and a clear need to improve multilingual communication, use of these tools in health and other critical real-world settings remains slow.

One important factor is that the varying accuracy of automated speech recognition and machine translation requires human correction—especially given the wide range of speakers, accents, recording quality, and other factors (Dew et al., 2018). Another challenge, particularly for HNA, is that in some countries currently affected by humanitarian crises, commercial tools for ASR and MT do not support the local languages, including Amharic, Hausa, Rohingya, or Fulfulde (Abbott & Martinus, 2018; Gu et al., 2018). Alternative methods exist today to create high quality ASR and MT models without the use of commercial providers. Examples include Coqui (2022) and Mozilla’s (2022b) DeepSpeech, two open-source libraries for ASR and MT

that can be used with freely available corpuses of language libraries, such as Mozilla's (2022a) CommonVoice and Tatoeba (2022). Such public repositories tend to favor languages spoken in wealthier countries, which drives organizations such as Translators Without Borders to build ASR and MT models specifically for languages spoken in developing countries experiencing some of the most protracted humanitarian crises (Ansari & Petras, 2018). Recent efforts by Google Research to close this gap have also resulted in 24 additional MT languages being added to "cover languages with large speaker populations in regions that are under-represented in technology" (Bapna et al., 2022, p. 5). Finally, depending on audio quality and the accuracy of the ASR model for a given language, errors can often be found in automated transcripts, which typically require human verification in order to ensure that the text is accurate (Basma et al., 2011).

The second of the three major challenges is that there are significant ethical issues that could stem from carelessly created language models for transcribing or translating human speech—as well as software that make these technologies more available—which could negatively affect vulnerable people affected by humanitarian crises. Ethical issues are defined here as actions that may not conform to relevant moral standards, particularly those moral standards implicit in relevant humanitarian principles because of the risks they present to affected populations or to organizations providing humanitarian assistance (Slim, 2015, pp. 47–145). Such risks are particularly acute when processing personal data, such as biometrics collected for refugee assistance or individuals' names that may be deemed necessary for delivering assistance (International Committee of the Red Cross and Privacy International, 2018; Raymond et al., 2016; Roberts & Faith, 2021; The Engine Room and Oxfam, 2018; UN Global

Pulse, 2016). In particular, NLP technologies have the potential to introduce new ethical issues in humanitarian assistance: Previous research has shown that some ASR technologies were less accurate when transcribing speech from women (Garnerin et al., 2019; Tatman, 2017) or from African Americans (Koenecke et al., 2020). Such issues can potentially introduce more errors in datasets—or mean that some groups are not represented at all and thereby result in what Sarı et al. (2021) termed “fairness gaps.” Without addressing such disparities when deploying NLP in practice, blind reliance on inaccurate transcription or translation models could in the future affect high-level decisions for distributing aid by magnifying gender, racial, or cultural biases.

Humanitarian organizations have increasingly become directly involved in the software innovation process (Betts & Bloom, 2014; Obrecht & Warner, 20116; Ramalingam et al., 2009), giving them the chance to design tools that best benefit people affected by humanitarian crises. Smith et al. (2020) propose that humanitarian organizations explicitly incorporate ethical values into their design process, for example by following practical guidelines that aim to mitigate ethical issues during the initial innovation stage (Elhra & Humanitarian Health Ethics, 2021; Krishnaraj et al., 2021). Many humanitarian organizations have established guidelines that aim to protect against ethical issues, particularly those that stem from processing personal data (IASC, 2021; International Committee of the Red Cross, 2020; OCHA, 2019), or that help their organization weigh the harms and benefits of adopting new innovations (Sheather et al., 2016). But because of the nature of software innovation, technical guidance documents can quickly become out of date—particularly as artificial intelligence systems become more widely used to assist with data processing.

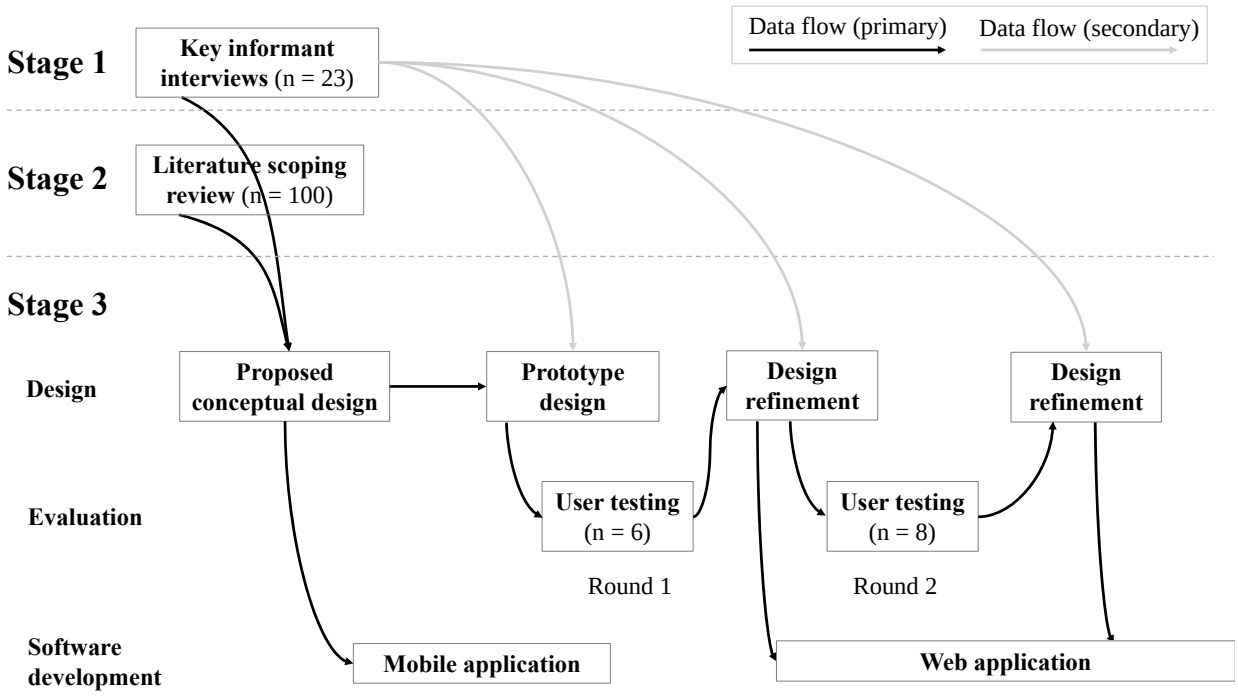
The third major challenge is that many software innovations suffer from such poor design that it can render them unusable, regardless of good intentions and substantial financial resources. Usability is understood here as the "extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use" (International Organization for Standardization, 2018). Usability is particularly problematic for publicly funded software projects that often fail to deliver the expected impact, despite considerable financial resources (Anthopoulos et al., 2016; Savoldelli et al., 2014). Best practices in web and software design therefore often follow the approach known as user-centered design (UCD), which is also known as human-centered design. UCD is understood here to be a framework for flexibly and iteratively creating innovations that are supported by knowledge about user realities, as well as knowledge of the institutional context in which software innovations are intended to be used (Holden & Boustani, 2021). UCD approaches typically require 1) a thorough understanding of the eventual users that takes their needs and environment into account; 2) empirical evaluations of user interactions, and 3) iterative cycles of designing and building new systems or services (Gould & Lewis, 1983). Failures during this process can mean that new software technology tools are rejected by organizations or professionals, or can have the potential to cause harm (Cornet et al., 2019). These challenges are unique and acute in critical settings such as during the provision of humanitarian assistance in humanitarian emergencies, or in providing healthcare to extremely vulnerable populations by often overstretched professionals (Ben-Zeev et al., 2015; Kushniruk & Nøhr, 2016). Similarly, whereas UCD is often embraced in theory, evaluating new tools

empirically based on user testing has many challenges that can dissuade humanitarian innovators from embracing this practice more widely (Cornet et al., 2020).

Given the three major challenges described above, the aim of this study was to investigate ways in which new features for the KoboToolbox software might be designed, to provide specialists involved in managing HNA data with the ability to systematically transcribe, translate, and analyze answers to OEQ from HNA interviews. This study formed part of a broader mixed-methods program of research to design tools to improve the quality of HNA by exploring the potential role played by different NLP technologies, as described in Chapter 3 (Kreutzer et al., 2020). The study was based on a three-stage, mixed-methods approach that included qualitative interviews, a scoping review, as well as an iterative user centered design process accompanied with usability testing, as diagrammed in Figure 9. In Stage 1, data from interviews with 23 key informants (KIs) involved in conducting HNA about proposed technological innovations were analyzed with the goal of informing the functional requirements of the new features. Stage 2 identified relevant details from a scoping review of ethical issues and their implications, stemming from the processing of data relating to people affected by a humanitarian crisis, in order to inform ethical considerations during the design process. Results from stages 1 and 2 are described in full detail in Chapters 4 and 5 (Kreutzer, 2021b, 2021a). In Stage 3, data about the iterative design process of the new software features, as well as usability testing results based on observations and feedback from 14 test users during the design process, were acquired and analyzed.

Figure 9

Overview of Methods and Data Collected for this Study. Graphic based on Cornet et al. (2020)



6.1. Methods

This study followed a mixed-methods Design Science Research (DSR) approach, whereby design is considered "disciplined, reflective discovery, and development of concepts for seeing what is possible and methods for realizing what is possible " (Jackson & Aakhus, 2014, p. 3). Whereas user-centered design is one particular method for creating highly usable software, DSR is a widely accepted research method for understanding the process of designing new solutions (Adikari et al., 2009).The DSR cycle typically starts with an awareness of a problem and a suggestion for a solution, resulting in a proposal and a tentative design. The further design and development of the new solution then leads to various forms of evaluation, leading to

iterative changes to the initial problem, further suggested solutions, and new development—until a conclusion can be reached at the end of the research project (Vaishnavi & Kuechler, 2015).

6.1.1. Stage 1: Key Informant Interviews

During Stage 1, 23 KIs were selected from 13 countries, representing 17 humanitarian organizations. A semi-structured interview guide was developed to: 1) investigate the contextual considerations for using quantitative and qualitative methods in HNA; 2) to understand the constraints related to qualitative interview methods, and 3) to solicit proposed concrete solutions for improving the relevance of qualitative interview methods used to gather data to inform humanitarian assistance. Data collection took place between November 2020 and May 2021. Full methodological details are provided in Chapter 4 (Kreutzer, 2021b).

6.1.2. Stage 2: Literature Scoping Review

A scoping review was conducted to map the range of ethical issues that have been raised in the academic literature regarding the processing of data from people affected by humanitarian crises. The scoping review using the method established by Arksey and O'Malley (2005), and further refined by Levac et al. (2010), following the framework maintained by the Joanna Briggs Institute (Peters et al., 2015, 2017). Comprehensive literature searches of 20 databases were conducted using Ovid, Ebsco, Web of Science, and Proquest and based on a list of iteratively developed keywords to identify peer-reviewed studies published since 2010. Further methodological details are available in Chapter 5 (Kreutzer, 2021a).

6.1.3. Stage 3: User Interface Design, Testing and Software Development

Design. Based on the needs identified in Stages 1 and 2, a proposed conceptual design of new features for the KoboToolbox software was developed by the researcher. The proposed

conceptual design was then used to inform an iterative creation process of comprehensive interactive prototypes by a user interface design specialist at KoboToolbox, using the Figma user interface design software (Figma Inc., 2021). The user-centered design process followed the approach proposed by Gould & Lewis (1983) as described above. The prototypes were intended for use on desktop computers where they could be accessed through the Figma web application, allowing users to interact with it. After the initial prototype design stage, designs were refined following each round of user testing, as shown in Figure 9.

User Testing. The research identified 14 user testing participants who possessed prior experience working with qualitative data, either in the role of managing data collection, transcribing audio responses, or analyzing collected data. They took part in two testing sessions, intended to test the usability of the new KoboToolbox software features by systematically recording observations made by the researcher about the test users, as well as and oral feedback received from the participants (Travis & Hodgson, 2019). Only participants who spoke either English, French, German, or Spanish were considered as these were languages spoken by the researcher. Participants were chosen using purposive sampling from a list of volunteers. Initial volunteer recruitment was done by posting a message on a web-based community forum to ask for volunteer participants with relevant experience. For Round 1, participants were also sought specifically at an organization that provides translation services to many humanitarian organizations using qualitative and recorded verbal interview methods. For Round 1, six participants were recruited. For Round 2, eight participants were recruited. Participants were sought to maximize heterogeneity based on geographic location, gender, type of organization, and level of experience working with qualitative data. In addition, participants in Round 2 were

also identified based on whether they spoke at least one additional language other than English in order to be able to transcribe and translate audio responses. Identified participants were informed about the study by email and informed consent was obtained from them. The consent form indicated that a screen recording of the user test might be created (if explicit consent was given) but that all responses would be treated confidentially.

Data collection took place in October 2021 and in July 2022 for the first and second rounds of user testing, respectively. User testing was conducted remotely using the Zoom platform (Gray et al., 2020), while video recordings of users' screens were created to support analysis. Participants were given access to the design prototypes during the course of the first testing session, and access to the actual software on a staging server for the second round. During each testing session, participants were asked to look at various prototype designs and software features on their computer and interact with them. They were given simple prompts to execute several tasks, such as "You would like to translate this transcript from Swahili to English using the automated option. How would you do this?" Participants were told that they could ask questions about functionality and were asked to "think out loud" about their use of the prototypes throughout each task. In all rounds, participants were asked a series of open-ended questions after completion of the exercises, to gather general feedback and suggestions, and to find out whether they found the new functionalities to be intuitive, which aspects they struggled with or would modify, and whether or how they would use the new features in their own work. The list of tasks and prompts and questions for each round of user testing can be found in Chapter 6 - Appendix 1.

Detailed anonymized notes were taken for each user testing session by the researcher. Notes for observations about how each task was performed by participants, as well as a summary of oral responses given to each open-ended debriefing question, were entered into a spreadsheet during the user testing session and were further expanded or corrected based on the recording, as needed. A summary of notes was created after each of the two rounds of user testing to extract relevant comments and individual recommendations to inform necessary modifications to the new KoboToolbox software features .

Software Development. Software engineers implemented the new software features using agile software development methods (Holden et al., 2021). Software development began after the completion of Round 1 of user testing based on the first design refinement, as shown in Figure 9. New features were then deployed to a staging server for Round 2 of user testing. Software code was updated after the second design refinement to fix pressing issues based on observations and feedback received from participants from Round 2.

6.2. Results

6.2.1. Stage 1: Key Informant Interviews

23 KIs (61% female) residing in 13 countries from 17 humanitarian organizations participated. The mean duration of interviews was 89 minutes (SD = 26; range = 46–144 min).

Analysis of the transcribed interviews resulted in 47 themes. Full results and discussion of the qualitative interviews, particularly around the contextual considerations, limitations, and non-technological proposals, are published in Chapter 4 (Kreutzer, 2021b).

Of the 47 themes, 25 themes were assessed qualitatively to identify issues that related to or could be addressed through technological innovations in a CAPI tool, as shown in Table 19

and summarized in this section, including quotes from KIs that are representative of similar responses.

Table 19

Themes Derived from KI Interviews Related to Technological Innovations in CAPI Tools

	Theme
Prevailing Approaches to Data Collection in HNA	
1	The prevalence of quantitative data has increased due to the widespread usage of KoboToolbox
2	OEQ are considered crucial for understanding complex subjects, but answers are mostly grouped as pre-coded options
3	Uncoded answers to OEQ can be crucial for filling in information gaps
4	Paper and Excel are often preferred over CAPI tools for taking long notes in KII
5	Qualitative methods are critical for answering “why” and “how” questions
6	Qualitative and quantitative methods are seen as complimentary
Constraints related to quantitative interview methods	
7	Some information by respondents is never recorded
Constraints related to qualitative interview methods	
8	Qualitative analysis is too time consuming for emergency contexts
9	There are not enough staff with expertise in qualitative methods
10	Interviewer training is much too short
11	Manually created notes are often of poor quality
12	Verbatim transcripts of interview responses are only rarely created because audio is almost never recorded
13	Audio files can create data protection risks or make respondents more reticent
14	Notes are rarely translated properly
15	Meaning is lost because analysis is not done by or with local staff
16	Qualitative and quantitative methods are often not integrated
Challenges specific to pre-coded open-ended questions	
17	Pre-coded responses to open-ended questions in questionnaires can lead to false data
18	OEQ are considered an obstacle to obtaining large samples
19	Most text responses to OEQ are not properly analyzed
20	Interviewers may not choose the correct responses among pre-coded options
Suggestions for improving the role of qualitative interview methods in HNA	
21	Conduct hybrid surveys with a mix of quantitative and qualitative methods
22	More audio recordings would be useful to capture all details
23	Automatic transcription would allow more widespread qualitative data collection
24	Machine translation would be helpful for systematically translating transcripts or notes
25	Automatic coding would be helpful to speed up analysis

Interviews with KIs showed a strong need for qualitative data for HNA in many situations, including a desire to use mixed-methods approaches and include OEQ in surveys:

> When we don't use qualitative data, we might really miss out on really key elements. If we miss out on having the quantitative component, we also miss out on representativity. . . . So, in an ideal case scenario, we have both. (KI #13)

Many KIs argued that OEQ, and qualitative methods in general, are crucial for understanding complex subjects related to the needs of affected populations, because "people might speak more to their unique situation." (KI #03) Some were very critical of the view that "the ideal needs assessment [at the UN level] is a household survey" (KI #05), instead arguing for more qualitative methods to provide more accurate information.

For most KIs, existing tools do not easily lend themselves towards collecting more qualitative information during HNA. KoboToolbox was brought up by all respondents as a tool that enabled rapid quantitative data collection:

> Kobo's been set up in such a way where it's so easy to get large volumes of data. . . . That day, you can do a survey and, in the evening, get back—or even in the car on the way back from the site, you can start to look at the data almost immediately. It's amazing. (KI #08)

However, KIs also pointed out that using CAPI tools like KoboToolbox were not conducive to typing in notes to OEQ due to the small screen size (they often preferred paper or laptop computers), whereas recording responses as an audio file during CAPI interviews was rarely done. KIs also reported that limited budgets or available time also meant that OEQ were in

some cases taken out of household surveys to “get a better picture of more people” (KI #03) through bigger samples. KIs cited a significant lack of enumerators or analysts trained in qualitative methods, making qualitative assessments often dependent on small teams of “people that are coming with quite a lot of experience” (KI #12) who are flown in during an emergency phase. KIs reported that the role of international staff often required notes or transcripts to be translated into English or another language spoken by the international team—a practice one KI criticized passionately:

> Why don’t we show the qualitative data so that it must be in the native language? . . . That’s not good enough. There are so many words that change the entire meaning, or it will not exist in a certain vocabulary, that can completely screw your assessment. (KI #21)

Some interviewees noted that even when notes were collected during interviews, “most of the time, [the notes to OEQ] don’t get translated” (KI #05) and are therefore not considered for analysis. Systematic time and funding shortages were seen as a barrier to addressing this gap: KIs noted that enumerator training was often limited to a single day, which was already considered an improvement as “there have been cases in the past where we would do like a one-hour training, and that that is the reality of NGOs sometimes” (KI #14).

Respondents provided several suggestions related to innovations in CAPI tools that would help overcome these constraints. Some suggested creating a mixed-methods hybrid survey that replaced many closed-ended questions with open-ended ones for which audio is recorded and then later analyzed into separate variables, arguing “we could then have far more representative data about a population at a much more intimate and deeper level” (KI #03). Some

KIs suggested that this would require fewer closed-ended questions, with more variables to be created during the analysis stage:

> I could probably cut my survey in half. “Which are your priorities right now?”, and then “This, this, and this”. “Okay, why? Tell me.” And then I can analyze the why. I could get rid of at least 20-30 questions, easily. (KI #21)

Finally, many KIs said that having an easier method to record, store, and process audio recordings of interviews would be beneficial: “If I could choose, I would always go for transcripts and recordings” (KI #01). Several KIs recommended leveraging automated means for processing responses to OEQ, including MT and automated coding:

> If . . . initial coding or sorting of that data could be done in an automated, quick way . . . that would change everything. (KI #02)

6.2.2. Stage 2: Literature Scoping Review

The database literature search returned 8,387 citations, of which 100 were included for data extraction following all screening stages.

The scoping review identified 22 ethical issues, which are described in more detail in Chapter 5 (Kreutzer, 2021a). Of the 22 ethical issues, seven were identified qualitatively as having the potential to be addressed through technological innovations in CAPI tools, as opposed to organizational measures or procedures. The seven issues, as well as the number of studies in which they were identified, are displayed in Table 20.

Table 20

List of Ethical Issues Identified Related to Technological Innovations in CAPI Tools

	Ethical Issues	Number of Studies
1	Unreliability: Processed data are inaccurate and do not sufficiently reflect reality to inform assistance	61
2	Lack of consent: Data are collected without informed consent	52
3	Dependence: Data are processed with the assistance of a political, economic, or military entity	48
4	Data security: Personal/sensitive data are not protected against malicious actors	42
5	Lack of respect: People/communities are not treated with respect	30
6	Excess: More data than necessary were collected	17
7	Undisclosed use: Data may be used beyond purposes for which they were collected	8

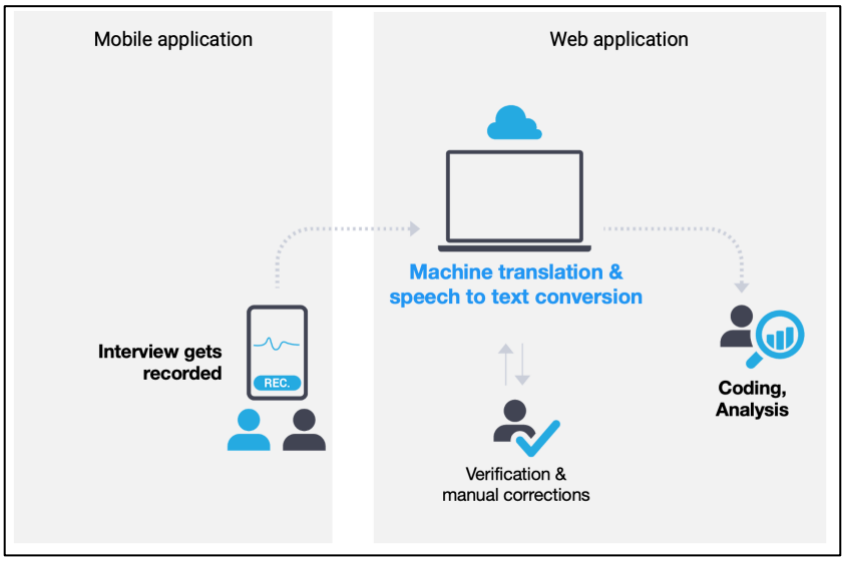
The most frequently cited ethical issues were data being collected in ways that are unreliable so that data is not sufficiently representative of reality to justify assistance (n = 61) and a lack of informed consent about how and for what purposes data were being processed (n = 52). For example, Shoemaker et al. (2019) found that refugees were frequently asked by humanitarian organizations to provide personal information that these refugees considered intrusive, without being told why this was relevant—and feeling that they lacked agency to decline. Other issues included a dependence on political, economic, or military entities for processing data (n = 48). Often, this was related to unclear data processing by commercial providers, insufficient data protection against malicious actors (n = 42), a lack of respect in connection with processing data (n = 30), collecting more data than needed (n = 17), and secondary uses of data beyond the initial purpose (n = 8).

6.2.3. Stage 3: User Interface Design, Testing and Software Development

Design. Relevant themes from Stage 1 (as displayed in Table 1) as well as relevant ethical issues from Stage 2 (as displayed in Table 20) were used to inform the proposed conceptual design. The proposed conceptual design is diagrammed in Figure 10 and included in detail in Chapter 6 - Appendix 2. It described at a high functional level how new features in KoboToolbox would allow users to 1) better collect audio responses during HNA interviews, 2) create transcriptions and translations of the audio files using NLP methods, 3) make corrections to the NLP-generated text as needed, and 4) review and download the data. The proposed conceptual design required extending both the main web application of KoboToolbox and its mobile data collection application, KoboCollect (KoboToolbox, 2022a)

Figure 10

Diagram of Proposed Conceptual Design



Ethical issues and needs identified by KIs were linked to specific decisions about the need for particular features. In some cases, design suggestions were identified that addressed

different and multiple themes or ethical issues. The design decisions, as well as the relevant ethical issues or themes that informed them, are listed in detail in Table 21.

Table 21

Design Decisions Linked to Themes from Key Informant Interviews (Stage 1) and Ethical Issues Related to CAPI Software (Stage 2)

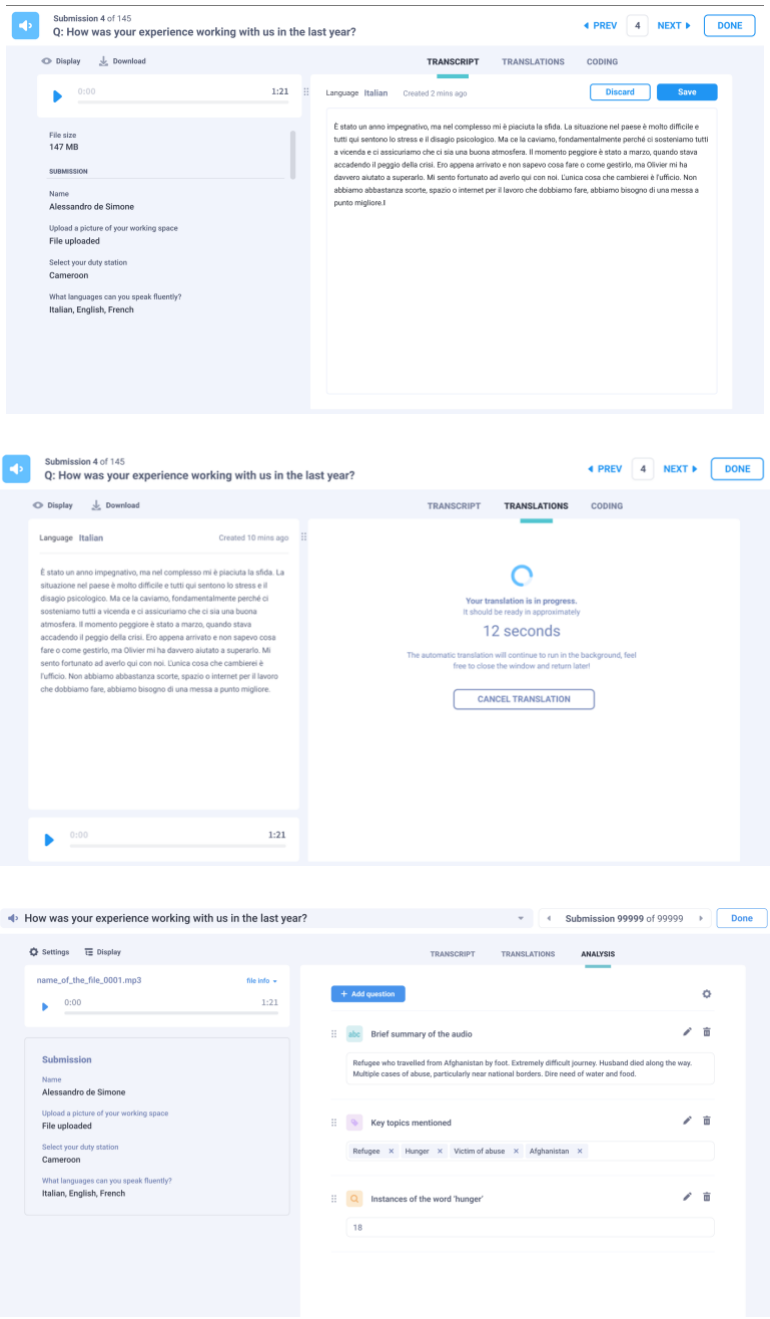
Feature	Design Decision	KII Themes	Ethical Issues
1. Justification for innovation	Processing qualitative data should be as intuitive, user friendly, and as easy to train as possible	9, 8, 10, 15	
	Integrate more qualitative features to be used alongside quantitative CAPI features in KoboToolbox	1, 6,16, 18, 20, 21	
	Create features for analyzing uncoded OEQ qualitatively post-data collection	2, 3, 5, 17	5
2. Data security	Ensure high level of data security; educate users about proper data protection measures	13	4
	Allow deletion of transcripts, translations, or audio recordings		6
	Enable removal of personal information; allow deletion of audio recordings		
3. Transcription	Allow automated/bulk processing of audio recordings	18	
	Enable rapid human verification of transcripts		
	Users should be able to turn audio recordings into more complete notes as part of the analysis process	11	
	Allow easy control over audio playback while creating notes or transcripts	4	
	Users should be able to rapidly create a transcript using ASR	23	
	Manual transcription should not require additional set up or training		
4. Translation	Enable users to translate manually created notes	14	
	Allow manual or machine translation of notes	24	
	Enable users to verify MT text		
5. Open-source ASR and MT	Do not allow secondary processing, e.g., for improving ASR or MT algorithms without consent		7
	Create a method for ASR and MT that does not require the use of commercial services		3
6. Qualitative analysis	Enable users to do rapid categorization of responses to OEQ	19	
	Ensure that categories can be updated and amended easily		

Feature	Design Decision	KII Themes	Ethical Issues
	Explore the use of automated content analysis tools (future enhancements)	25	
7. Mobile application	Add warning for background recording in mobile application when the form is opened		2
	Propose requesting consent as part of form		
	Users should be able to record audio in the background for entire interview	7	1
	Users should be able to create specific audio questions alongside one or more regular questions	12	
	Allow users to compare audio with structured data for accuracy	22	

The comprehensive interactive prototypes for the web application were developed iteratively. Bi-weekly meetings between the user interface design specialist and the researcher were held to discuss progress and to make decisions about potential solutions based on the data collected during Stages 1 and 2. Initial designs focused on processing audio files for ASR and MT in bulk, for example, to transcribe all audio files at once. This option assumed a completely separate view for reviewing and processing audio files. However, based on a review of results from Stage 1, a decision was made to focus primarily on processing individual audio files first, given the importance of making manual corrections to automatically generated transcripts or translations, as well as the need to create these texts manually. This decision also signaled the need for additional designs to modify the existing data table feature to allow users to access new transcription, translation, and analysis features for processing and viewing OEQ data directly in the existing data table view—a complex design requirement given the limited amount of space available in the user interface. Additional designs covered new functionality for exporting transcript data for use in external applications, as well as different navigation options to switch between respondents and between different audio responses from the same respondent.

Figure 11

Examples of the Prototype Designs used in Round 1 of User Testing



Interactive prototypes related to the transcription and translation features were created for the purpose of conducting the first round of user testing. Examples of designs used for Round 1 of user testing can be found in Figure 11.

User Testing Round 1. A total of six candidates (two of whom were female) from four countries participated in the first round of user testing (three from Nigeria, one each from Argentina, Cambodia, and Switzerland). The mean duration of user testing sessions was 46 minutes (SD = 6.4; range = 39–53 min). Three participants were active KoboToolbox users with significant qualitative data collection experience, whereas three participants had little or no previous experience using KoboToolbox but were involved in transcribing and translating interview responses with other tools. One interview was conducted in Spanish with the remainder conducted in English.

Most participants were able to do all the assigned tasks with no or minimal prompting. All participants evaluated the new features as positive and useful for their work. Evaluations with test users pointed to eight usability issues (one for the table view, three for the transcription feature, three for translations, one related to the analysis feature, and one for navigation between audio questions), as shown in Table 22.

For each item, a specific design change was identified that would address the issue. For example, three testers did not recognize the “Process” button as the way to access transcription and translation features and required additional prompting, which was changed to show an obvious way of playing the audio in the table view, in addition to a button labeled “Analyze”. Two participants found a list of technical options for conducting ASR confusing, which was addressed by using a single default ASR and MT method. Another participant argued in favor of

allowing users to navigate between responses to the same question, as well as between responses to different qualitative questions from the same respondent, which was addressed by creating more obvious way of switching between questions and moving to the next or previous interviewee.

Specific design changes were implemented based on the findings from the results of user testing round 1. The updated set of designs was then used as the basis for implementing the new web application software features in KoboToolbox.

Table 22

Overview of Issues from User Testing Round 1 and Decisions on Design Changes

Feature	Feedback	No. of participants	Design change
Table	Did not recognize "Process" button for accessing transcription/translation feature	3	Rename or make button more intuitive
Transcription	Expected the question above the transcript, and the respondent name/ID at the top (where the question currently is)	1	Make the labeling more intuitive Move text editing section under the question label
	Was not sure what different ASR providers meant	2	Choose a provider by default, allowing users to edit if needed
	Did not immediately understand difference between automatic and manual transcription	2	Add explanation labels
Translation	Did not find the "Begin" button for creating first translation	1	Rather than "Begin", name the button "Create transcript" and "Create translation"
	Did not see language variant dropdown for UK after selecting English	1	Make language variant dropdown more intuitive
	Ability to compare any translation or transcript side by side	1	Add dropdown to switch text in left panel
Analysis	Was confused by the Coding tab, assumed it referred to adding timestamps rather than analyzing qualitative data	1	Rename from Coding to Analysis
Navigation	Did not find an easy way to switch to next respondent in the same question or next audio question for same respondent	1	Add additional navigation options

Other suggestions were considered useful for further development of the new features but were not possible to implement given time and budget constraints (see Table 23). This included ideas such as differentiating speakers in focus groups or other group interviews, adding timestamps to long responses, or marking the quality of audio or generated text, transcribing directly from one language into another, or seeing visually whether a translation was done manually or through MT.

All participants stated that the new features would be useful in their work, particularly for saving time. Four participants liked that the user interface was intuitive in their opinion, two of whom suggested that this would make onboarding of their teams easier.

Table 23

Proposed Additional Features Identified During User Testing

Feature	Feedback	No. of Participants
Text editing	Offline functionality: Not losing work at any point in time, maybe display warning when connectivity is lost or unstable.	1
	Basic text styling functionalities to emphasize parts of the text, different speakers, etc.	3
	Spell Check functionality (or integration with existing dictionary)	2
Transcription	Automatically recognize different speakers	1
	Add a rating for audio quality 1-5. Helps knowing which audios to prioritize.	1
	Include timestamp functionality	3
	Keyboard shortcuts for pausing audio, slowing down, speed up, etc.	1
Translation/ transcription	One respondent at a time: See more audio questions from the same respondent.	1
	Audit functionality/history of changes to the text (transcript and translation)	2
	Translations/transcription verification. Checkbox where each can be approved.	3
Analysis	Tag/add metadata at the sentence level and the respondent level	1
	Functionality to do some pattern recognition	1
General	Indication of whether the translation was done manually or by MT, or reviewed already.	1
	Ability to transcribe and translate directly from the source to the target language	2

User Testing Round 2. A total of eight candidates (three of whom were female) from eight countries participated in the second round of user testing (one each from Afghanistan, France, Guatemala, India, Indonesia, Mexico, Nepal, and Nigeria). The mean duration of user testing sessions was 47 minutes (SD = 10; range = 35–67 min). Four participants worked for international organizations, two for an academic institution, and one for a local non-governmental organization. All participants were active KoboToolbox users of whom three reported to collect qualitative data somewhat often and five said they were frequently involved in

collecting qualitative data. One interview was conducted in Spanish with the remaining user testing sessions conducted in English.

Tasks given to participants involved transcribing and translating interview responses from audio files. All respondents were asked to transcribe files in English as well as in another language spoken by the participant. Among them, participants transcribed into and translated from audio files in Dari, English, French, Haussa, Hindi, Indonesian, Nepali, and Spanish. All candidates were able to complete the list of tasks with no or minimal prompting. Changes implemented based on the results from Round 1, as shown in Table 22, were confirmed to have resolved the earlier usability issues. A set of smaller new usability issues were identified based on observations and user feedback with corresponding design changes identified, which are shown in Table 24.

Candidates were first asked to transcribe audio to text. Audio files provided were intentionally of low to medium quality in order to make the transcription more challenging and to require manual corrections in case ASR was chosen. In all cases and without being prompted or explained what it refers to, they chose “automatic” transcription rather than the manual option. When asked to explain the functionality, users instinctively described correctly how ASR worked, with one explaining it meant “the machine will do it for me” (P4). Most participants judged the ASR-generated transcription to be good. Most users only found errors in the text, either immediately or after being prompted by the researcher. Quality of ASR text appeared to depend on audio quality rather than languages, with one respondent judging the ASR result from English to be much poorer than that from an Indonesian response, even though both were recorded from the same respondent. Participants were found to correct transcripts quickly by re-

listening to audio files and pausing to make corrections. Several participants did not correct all mistakes in transcripts from English recordings, possibly because English was not the first language for any candidate. Conversely, corrections in transcriptions in candidates' native language appeared to be easier, with participants appearing to correct transcripts more from memory than from re-listening to specific parts of the audio file.

All participants immediately found the translation feature without requiring help. When searching for the target language, one user typed in "Español" and did not get a result in the language dropdown, but instinctively searched for "Spanish" instead. Similarly, one user trying to translate from English into Dari (the main language spoken in Afghanistan) found that it was not available for automatic translation and instead searched for "Persian" (which overlaps significantly with Dari), for which they found MT was available. Similar to the transcription exercise, all participants chose to use MT as the translation option. When asked to judge the accuracy, participants found fewer issues and corrected them more quickly compared to transcripts. Although the original transcript is shown alongside the translation, one user translating from Hindi said they preferred listening to the original audio in the source language to verify if the English translation was correct, rather than comparing it to the English transcript.

The most common UI challenges identified were participants struggling to see an overview of all created transcripts in the table (observed with four participants), trying to switch between audio questions from the same respondent, and expecting transcripts to be downloaded from the single submission view rather than from the Downloads section (observed with three participants, respectively). The first two were addressed through a software change whereas the

second was remedied through a small hint in the UI that points users to the correct screen for downloading transcript data, as summarized in Table 24.

Several suggestions were noted that may be used in future enhancements of the new features, but which were out of scope for the present iterations. One user suggested adding a visual cue for users to encourage them to learn more about on ethical ways of collecting and processing audio data. Other suggestions included a way to store “favorite” languages, a new permission level for only editing transcripts and translations, see a correction log of changes made by different users, and adding timestamps to help with very long audio.

As in the first round, participants found the transcription feature to be very useful for their work to save time (mentioned by five participants) or to enable them to collect data more efficiently by replacing other kinds of questions with an audio question instead (mentioned by three participants).

Table 24

Overview of Issues from User Testing Round 2 and Decisions on Design Changes

Feature	Feedback	No. of participants	Design change
Table	Tried to transcribe audio in the submission edit feature	2	Rename button to “Transcribe”
	Reviewing transcripts in table view was not always easy because of auto-generated empty columns for each ASR/MT language	2	Hide empty ASR/MT columns; remove “N/A” from empty cells
Transcription	Was unclear about what ‘manual’ and ‘automatic’ buttons referred to	1	Provide explanation text in the UI
	Did not understand that ASR text was in edit mode / had to be saved before being able to continue	1	Provide guidance in the UI
Translation	Didn’t see that original transcript text was displayed on the right while reviewing MT text	1	Make the transcript field more obvious during translation
Navigation	Couldn’t switch to different audio question of the same respondent; when choosing different question the first respondent would be displayed	3	Change how question navigation works
Downloads	Expected transcript download to be in single submission view rather than Downloads section	2	Provide guidance in the UI

Software Development. The software code for new features created for the web application of KoboToolbox were released on GitHub (KoboToolbox, 2022b). New features are available for humanitarian and other organizations to use on hosted cloud servers (Kobo, 2022).

Specific changes to the web application software involved the following items:

1. Building a new user interface for creating, displaying, and interacting with transcription data;
2. Adapting the database model in order to store transcription, translation, and analysis related to particular fields in the dataset;
3. Creating new application programming interface (API) endpoints for accessing these fields;

4. Creating a method for hosting and integrating with open-source ASR models to provide automated transcription for languages that are not yet available commercially;
5. Creating integrations with commercial ASR and MT providers; and
6. Expanding data export features to be able to customize how new qualitative data can be downloaded to other tools.

In terms of NLP technology, the new software features integrate ASR and MT capability provided by Google Cloud Compute, which as of 2022 provide 72 languages for transcription (including 138 regional variants) and 106 languages for translation. Users can also choose from a list of 6,092 languages for doing manual transcription or translation. This list was established by the researcher using the ISO 639-3 (2007) comprehensive list of languages, made available by SIL International (Eberhard et al., 2022), and then filtering for any living languages and excluding sign languages. For any language that has ASR or MT, users are able to choose between “manual” and “automatic” transcription or translation; for the remaining languages only the manual method is available.

Changes to the KoboCollect mobile application were implemented to allow easier recording of audio without the need of additional applications. Specifications, user requirements, and relevant implications for changes to related technologies were proposed, discussed, and decided upon through public consultations in an online forum (Kreutzer, 2019). Specifications were decided also for XLSForm (XLSForm.org, n.d.), a method used by KoboToolbox and other tools for creating complex CAPI questionnaires. The new feature allows users to set the audio recording quality through three pre-defined settings (voice-only, low, and normal, as shown in

Table 25). The first two settings were chosen to accommodate users working in low-bandwidth environments—a context that applies to many humanitarian crises—whereas the third was selected as providing the best balance between quality and file size. New audio recording features in KoboCollect, as shown in Figure 12, were released in April 2021 (KoboToolbox, 2022a).

Figure 12

Screenshot of Internal Audio Recording Feature Added to the Mobile Application

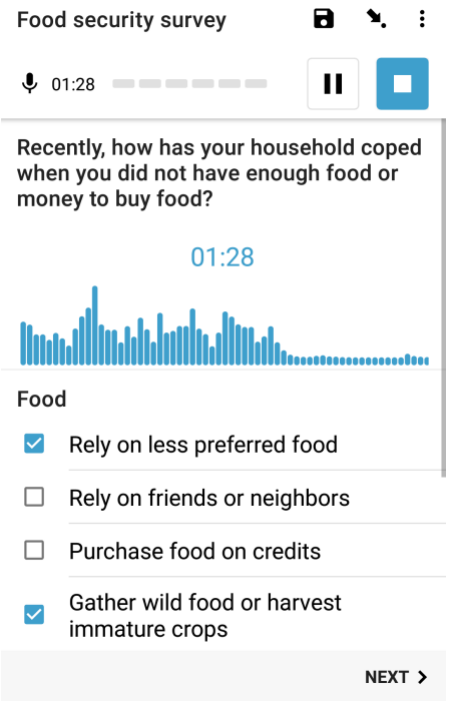


Table 25

Available Recording Settings to Accommodate Low-Bandwidth Users

Value	Extension	Encoding	Bit rate	Sample rate	File size
normal	.m4a	AAC	64kbps	32kHz	~30MB/hour
Low	.m4a	AAC	24kbps	32kHz	~11MB/hour
voice-only	.amr	AMR	12.2kbps	8kHz	~5MB/hour

6.3. Discussion

The aim of this study was to investigate how new features in the KoboToolbox software can be designed using user centered design principles to enable specialists involved in managing HNA data to systematically transcribe, translate, and analyze answers to open-ended questions OEQ from HNA interviews. This study formed part of a broader program of research into design tools that would improve the use of qualitative methods in HNA by exploring the potential role played by different NLP technologies, as described in Chapter 3 (Kreutzer et al., 2020). This study was based on a three-stage mixed-methods approach that included qualitative interviews (Stage 1), a scoping review (Stage 2), as well as an iterative user centered design process validated through two rounds of usability testing (Stage 3). This section describes the implications of the findings presented above, and then discusses the strengths and limitations of this study.

The user centered design process in Stage 3 was primarily shaped by the themes and ethical issues identified during Stages 1 and 2, as well as feedback from two rounds of user testing, and the professional experience of the researcher and user interface design specialist. User testing with 14 potential users showed strong support for the new features, particularly on the assumption that the new components would be easy to use and save time compared with

current methods of transcribing and translating audio recordings of responses to OEQ. User testing observations also confirmed that the focus on usability led to the intuitive use of the new features, especially for participants without any prior experience using KoboToolbox. The first round of user testing identified usability issues that needed to be addressed, such as non-intuitive buttons or navigation options. These issues were addressed in an updated design and reflected in the software that was built and made available on a public staging server. The second round of user testing, focused on using the actual software in a simulated survey setting for transcribing and translating audio recordings, confirmed that the first set of issues were addressed properly but also identified a smaller set of usability issues related to reviewing and downloading transcripts in bulk. Feedback from participants demonstrated that both automated and manual transcriptions are expected to make collecting qualitative interview responses faster than current methods. The findings also showed that users may be more likely to collect audio recordings to responses to OEQ as a way to replace existing questions. This idea, first proposed by KIs during Stage 1, recurred unprompted by different test users, thereby confirming the potentially widespread adoption of hybrid mixed methods surveys in the future.

The technical innovations included in the design of new KoboToolbox features were directly informed by the findings from Stage 1, which involved in-depth interviews with 23 purposively sampled KIs involved in conducting HNA that resulted in 25 themes that could be addressed through CAPI innovations. Some respondents voiced regret that a lot of useful information might be given during quantitative assessments, but that the structured nature of surveys has nowhere to record what was actually said. Instead, KIs generally expressed the need to be able to collect more qualitative data to inform humanitarian operations—alongside

quantitative methods—in order to provide a more complete picture than what is feasible through pre-coded multiple-choice questions. Interviews showed that time and resource constraints make up the most significant hurdles to collecting and using more qualitative data in HNA. Because proper analysis of qualitative data often requires significant time, qualitative methods are often considered to be unsuitable in emergency contexts where results need to be produced quickly.

Many KIs specifically proposed NLP methods that would allow for wider use of qualitative findings, in order to complement quantitative data for the purpose of accurately assessing population needs. The results from interviews with KIs showed that any technological innovations should address extreme constraints in terms of the skills of teams collecting and analyzing qualitative data. For the design process, this meant that any new features had to be highly intuitive and user friendly to avoid the need for additional training. Especially in emergency settings, KIs argued that collecting and analyzing data needs to be as fast and seamless as possible. As described in Chapter 4 (Kreutzer, 2021b), these types of technological innovations should be made in concert with other changes suggested by KIs for improving the quality and relevance of HNA, namely 1) acknowledging limitations of quantitative data, 2) prioritizing qualitative interview and analysis skills, and 3) implementing methodological innovations.

The user centered design process was also directly informed by numerous ethical considerations related to processing audio files with potentially sensitive interview responses. For example, multiple KIs during Stage 1 expressed the importance of protecting the personal data of respondents in their work, particularly when handling audio recordings that could more easily allow identifying the respondent or reveal sensitive information. A subset of seven ethical

issues (shown in Table 20) that could be addressed through software development were identified from the scoping review in Stage 2 that sought to map the range of ethical issues that have been raised in the academic literature regarding the processing of data from people affected by humanitarian crises. In Stage 3, each of the seven ethical issues was addressed in the design decisions for new features of the KoboToolbox CAPI software that was part of this study. For example, the ethical issue of being dependent on an external entity was addressed by creating an alternative way of conducting ASR and MT with non-commercial open-source methods. Results of the review showed a wide array of ethical issues that do not only require technological solutions, but especially organizational measures—as shown in more detail in Chapter 5 (Kreutzer, 2021a).

6.3.1. Strengths and Limitations

Strengths of this research include the following: a diverse sample of KIs with deep experience in HNA from an array of relevant humanitarian organizations and across a large global geography; a systematic approach to identifying ethical issues related to processing data in humanitarian assistance from all recently published academic studies; and a thoroughly-documented user centered design process that included two rounds of user testing with a diverse and international sample of potential users, to identify and address usability as well as ethical issues. This study was conducted in order to fill an important gap in academic research about the design of innovations for humanitarian assistance. This study has also contributed to the design science discipline by documenting the entire design cycle for new software features that can be used in many humanitarian crises over the coming years.

Limitations include relative under-sampling of KIs from Africa, East Asia, and Latin America as well as over-sampling of KIs from large international humanitarian organizations. For the scoping review in Stage 2 from which a subset of seven ethical issues was drawn, identifying relevant studies was a challenge due to the lack of a shared nomenclature across disciplines for humanitarian assistance, ethical issues, and data processing. Some potentially relevant publications that met the inclusion criteria may therefore have been missed. Although user testing candidates were chosen to represent the diverse community of potential users involved in qualitative data collection, it is impossible to ensure representativeness given the qualitative nature and small number of representatives in this kind of research. The two user testing rounds may also have missed divergent opinions and feedback from other types of users or organizations.

6.4. Conclusion

This study used a user-centered design and a mixed-methods approach to inform, design, and evaluate new software features for KoboToolbox. These new features have resulted in a set of practical tools that address professional requirements and can be readily deployed by humanitarian organizations to systematically transcribe, translate, and analyze answers to open-ended questions from HNA interviews. The needs of potential humanitarian users involved in HNA, as well as a subset of identified ethical issues related to using CAPI tools, were identified systematically through key informant interviews and a scoping review, respectively. The findings informed foundational decisions for the initial design of new KoboToolbox features. Two rounds of user testing identified several usability issues which were addressed during an

iterative user centered design process, but also showed strong support by potential humanitarian users for the applicability and usability of the new features in their work.

The multi-stage approach employed by this study has the potential to be used to guide other software innovations by humanitarian organizations to mitigate low rates of implementation, risks of some ethical issues, or costly modifications needed to fix usability issues later on. Further research is needed to evaluate the quality of ASR and MT based on audio recordings created in a humanitarian crisis setting and to measure the relative operational and informational benefits of using more OEQ during HNA.

6.5. References

Abbott, J. Z., & Martinus, L. (2018). Towards Neural Machine Translation for African Languages. *ArXiv:1811.05467 [Cs, Stat]*. https://arxiv.org/abs/1811.05467v1

Adikari, S., McDonald, C., & Campbell, J. (2009). Little design up-front: A design science approach to integrating usability into agile requirements engineering. *Human-Computer Interaction. New Trends. HCI 2009. Lecture Notes in Computer Science*, *5610*, 549–558. https://doi.org/10.1007/978-3-642-02574-7_62

Ansari, A., & Petras, R. (2018). *Gamayun: The Language Equality Initiative*. Translators Without Borders. https://translatorswithoutborders.org/wp-content/uploads/2018/03/Gamayun-Language-Equality-Initiative-March-2018.pdf

Anthopoulos, L., Reddick, C. G., Giannakidou, I., & Mavridis, N. (2016). Why e-government projects fail? An analysis of the Healthcare.gov website. *Government Information Quarterly*, *33*(1), 161–173. https://doi.org/10.1016/j.giq.2015.07.003

Arksey, H., & O'Malley, L. (2005). Scoping studies: Towards a methodological framework. *International Journal of Social Research Methodology*, *8*(1), 19–32. https://doi.org/10.1080/1364557032000119616

Assal, H., Seng, J., Kurfess, F., Schwarz, E., & Pohl, K. (2011). Semantically-enhanced information extraction. *2011 Aerospace Conference*, 1–14. https://doi.org/10.1109/AERO.2011.5747547

Banatvala, N. (2000). Conflict and health: Public health and humanitarian interventions: developing the evidence base. *BMJ*, *321*(7253), 101–105. https://doi.org/10.1136/bmj.321.7253.101

Bansal, S., Kamper, H., Livescu, K., Lopez, A., & Goldwater, S. (2018). Low-resource speech-to-text translation. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH, 2018-Septe*, 1298–1302. https://doi.org/10.21437/Interspeech.2018-1326

Bapna, A., Caswell, I., Kreutzer, J., Firat, O., van Esch, D., Siddhant, A., Niu, M., Baljekar, P., Garcia, X., Macherey, W., Breiner, T., Axelrod, V., Riesa, J., Cao, Y., Chen, M. X., Macherey, K., Krikun, M., Wang, P., Gutkin, A., … Hughes, M. (2022). Building machine translation systems for the next thousand languages. *ArXiv:2205.03983 [Cs]*. http://arxiv.org/abs/2205.03983

Basma, S., Lord, B., Jacks, L. M., Rizk, M., & Scaranelo, A. M. (2011). Error rates in breast imaging reports: Comparison of automatic speech recognition and dictation transcription. *American Journal of Roentgenology*, *197*(4), 923–927. https://doi.org/10.2214/AJR.11.6691

Ben-Zeev, D., Schueller, S. M., Begale, M., Duffecy, J., Kane, J. M., & Mohr, D. C. (2015). Strategies for mHealth Research: Lessons from 3 Mobile Intervention Studies. *Administration and Policy in Mental Health and Mental Health Services Research*, *42*(2), 157–167. https://doi.org/10.1007/s10488-014-0556-2

Betts, A., & Bloom, L. (2014). *Humanitarian Innovation: The State of the Art* (Issue November). UN OCHA. https://www.unocha.org/sites/unocha/files/Humanitarian Innovation The State of the Art_0.pdf

Building Markets, & Orange Door Research. (2018). *What is the Point if Nothing Changes: Current Practices and Future Opportunities to Improve Remote Monitoring and Evaluation in Syria*. New York, NY: Building Markets.

Coqui. (2022). *Coqui*. https://coqui.ai/

Cornet, V. P., Daley, C., Bolchini, D., Toscos, T., Mirro, M. J., & Holden, R. J. (2019). Patient-centered Design Grounded in User and Clinical Realities: Towards Valid Digital Health. *Proceedings of the International Symposium on Human Factors and Ergonomics in Health Care*, *8*(1), 100–104. https://doi.org/10.1177/2327857919081023

Cornet, V. P., Toscos, T., Bolchini, D., Ghahari, R. R., Ahmed, R., Daley, C., Mirro, M. J., & Holden, R. J. (2020). Untold stories in user-centered design of mobile health: Practical challenges and strategies learned from the design and evaluation of an app for older adults with heart failure. *JMIR MHealth and UHealth*, *8*(7), 1–14. https://doi.org/10.2196/17703

Cosgrave, J. (2009). Humanitarian funding and needs assessment. In *Humanitarian Response Index 2008: Donor Accountability in Humanitarian Action* (pp. 75–88). https://doi.org/10.1057/9780230584617

Dew, K. N., Turner, A. M., Choi, Y. K., Bosold, A., & Kirchhoff, K. (2018). Development of machine translation technology for assisting health communication: A systematic review. *Journal of Biomedical Informatics*, *85*(July), 56–67. https://doi.org/10.1016/j.jbi.2018.07.018

Eberhard, D. M., Simons, G. F., & Fennig, C. D. (Eds.). (2022). *Ethnologue: Languages of the World* (25th editi). SIL International. https://www.ethnologue.com/a

Elhra, & Humanitarian Health Ethics. (2021). *Ethics for humanitarian innovation: Toolkit*. https://higuide.elrha.org/wp-content/uploads/2021/09/Elrha-Ethics_for_Humanitarian_Innovation-Toolkit-210913.pdf

Emergency Data Science Workshop. (2018). *Natural Language Processing for Humanitarian*

Survey Work: Challenge Presented by International Rescue Committee. https://emergencydatascience.org/challenge_nlp_surveys/

Figma Inc. (2021). *Figma*. https://www.figma.com/

Galanis, N. I., Vafiadis, P., Mirzaev, K. G., & Papakostas, G. A. (2021). Machine Learning Meets Natural Language Processing - The Story so Far. *IFIP Advances in Information and Communication Technology*, *627*, 673–686. https://doi.org/10.1007/978-3-030-79150-6_53

Garnerin, M., Rossato, S., & Besacier, L. (2019). Gender representation in French broadcast corpora and its impact on ASR performance. *AI4TV 2019 - Proceedings of the 1st International Workshop on AI for Smart TV Content Production, Access and Delivery, Co-Located with MM 2019*, 3–9. https://doi.org/10.1145/3347449.3357480

Gerdin, M., Chataigner, P., Tax, L., Kubai, A., & von Schreeb, J. (2014). Does need matter? Needs assessments and decision-making among major humanitarian health agencies. *Disasters*, *38*(3), 451–464. https://doi.org/10.1111/disa.12065

Gould, J. D., & Lewis, C. (1983). Designing for usability-key principles and what designers think. *Conference on Human Factors in Computing Systems - Proceedings*, *28*(3), 50–53. https://doi.org/10.1145/800045.801579

Gray, L. M., Wong-Wylie, G., Rempel, G. R., & Cook, K. (2020). Expanding qualitative research interviewing strategies: Zoom video communications. *Qualitative Report*, *25*(5), 1292–1301. https://doi.org/10.46743/2160-3715/2020.4212

Gu, J., Wang, Y., Chen, Y., Cho, K., & Li, V. O. K. (2018). *Meta-Learning for Low-Resource Neural Machine Translation*. http://arxiv.org/abs/1808.08437

Hannun, A., Case, C., Casper, J., Catanzaro, B., Diamos, G., Elsen, E., Prenger, R., Satheesh, S.,

Sengupta, S., Coates, A., & Ng, A. Y. (2014). *Deep Speech: Scaling up end-to-end speech recognition*. http://arxiv.org/abs/1412.5567

Holden, R. J., & Boustani, M. A. (2021). Design and Usability Methods. In R. J. Holden & R. S. Valdez (Eds.), *The Patient Factor* (1st ed.). CRC Press.

Holden, R. J., Boustani, M. A., & Azar, J. (2021). Agile Innovation to transform healthcare: Innovating in complex adaptive systems is an everyday process, not a light bulb event. *BMJ Innovations*, *7*(2), 499–505. https://doi.org/10.1136/bmjinnov-2020-000574

IASC. (1999). *Protection of Internally Displaced Persons*. https://interagencystandingcommittee.org/system/files/legacy_files/protection_of_internally_displaced_persons_inter_agency_standing_committee_policy_paper_0.pdf

IASC. (2021). *Operational Guidance: Data Responsibility in Humanitarian Action*. Inter-Agency Standing Committee. https://interagencystandingcommittee.org/operational-response/iasc-operational-guidance-data-responsibility-humanitarian-action

International Committee of the Red Cross. (2020). *Handbook on Data Protection in Humanitarian Action (2nd Edition)*. https://www.icrc.org/en/data-protection-humanitarian-action-handbook

International Committee of the Red Cross and Privacy International. (2018). *The Humanitarian Metadata Problem: "Doing No Harm" in the Digital Era*. https://privacyinternational.org/sites/default/files/2018-12/The Humanitarian Metadata Problem - Doing No Harm in the Digital Era.pdf

International Organization for Standardization. (2007). *Codes for the representation of names of languages — Part 3: Alpha-3 code for comprehensive coverage of languages*. ISO 639-

3:2007.

International Organization for Standardization. (2018, March). *Ergonomics of human-system interaction — Part 11: Usability: Definitions and concepts*. ISO 9241-11:2018(En). https://kebs.isolutions.iso.org/obp/ui#iso:std:iso:9241:-11:ed-2:v1:en

Jackson, S., & Aakhus, M. (2014). Becoming More Reflective about the Role of Design in Communication. *Journal of Applied Communication Research*, *42*(2), 125–134. https://doi.org/10.1080/00909882.2014.882009

Klein, G., Kim, Y., Deng, Y., Senellart, J., & Rush, A. M. (2017). *OpenNMT: Open-Source Toolkit for Neural Machine Translation*. http://arxiv.org/abs/1701.02810

Kobo. (2022). *KoboToolbox* (2.021.45). https://www.kobotoolbox.org

KoboToolbox. (2022a). *kobotoolbox/collect*. https://github.com/kobotoolbox/collect

KoboToolbox. (2022b). *kobotoolbox/kpi*. https://github.com/kobotoolbox/kpi/

Koenecke, A., Nam, A., Lake, E., Nudell, J., Quartey, M., Mengesha, Z., Toups, C., Rickford, J. R., Jurafsky, D., & Goel, S. (2020). Racial disparities in automated speech recognition. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(14), 7684–7689. https://doi.org/10.1073/pnas.1915768117

Kreutzer, T. (2019). *Create automatic background recording of interview in Collect*. https://forum.getodk.org/t/create-automatic-background-recording-of-interview-in-collect/18601/

Kreutzer, T. (2021a). *Ethical Implications Related to Processing of Personal Data in Humanitarian Crises: A Scoping Review*.

Kreutzer, T. (2021b). *Overcoming Blind Spots: Constraints and Solutions Related to Qualitative*

Interview Methods in Humanitarian Needs Assessments.

Kreutzer, T., Vinck, P., Pham, P., An, A., Appel, L., Deluca, E., Tang, G., Alzghool, M., Hachhethu, K., Morris, B., Crowley, J., & Orbinski, J. (2020). Improving Humanitarian Needs Assessments through Natural Language Processing. *IBM Journal of Research and Development*, *64*(1). https://doi.org/10.1147/JRD.2019.2947014

Krishnaraj, G., Hunt, M., Schwartz, L., O'Mathúna, D., Pringle, J., Okhowat, A., Kapiriri, L., Brahimi, L., & Yantzi, R. (2021). *Ethics for Humanitarian Innovation: Background Paper*. Elrha. https://higuide.elrha.org/wp-content/uploads/2021/09/Elrha-Ethics_for_Humanitarian_Innovation-Background_Paper-210909.pdf

Kushniruk, A., & Nøhr, C. (2016). Participatory design, user involvement and health IT evaluation. *Evidence-Based Health Informatics: Promoting Safety and Efficiency through Scientific Methods and Ethical Policy*, *222*, 139–151. https://doi.org/10.3233/978-1-61499-635-4-139

Levac, D., Colquhoun, H., & O'Brien, K. K. (2010). Scoping studies: Advancing the methodology. *Implementation Science*, *5*(1). https://doi.org/10.1186/1748-5908-5-69

Mozilla. (2022a). *Common Voice*. https://commonvoice.mozilla.org/en

Mozilla. (2022b). *mozilla/DeepSpeech*. https://github.com/mozilla/DeepSpeech

Obrecht, A., & Warner, A. T. (20116). *More than just luck : Innovation in humanitarian action*. Humanitarian Innovation Fund and ALNAP.

OCHA. (2019). *Data responsibility guidelines*. UN Office for the Coordination of Humanitarian Affairs. https://centre.humdata.org/wp-content/uploads/2019/03/OCHA-DR-Guidelines-working-draft-032019.pdf

OCHA. (2021a). *Global Humanitarian Overview 2022*. Office for the Coordination of Humanitarian Affairs. https://doi.org/10.18356/9789210012423

OCHA. (2021b). *Joint intersectoral analysis framework 1.1*. UNOCHA. https://doi.org/10.18356/9789214030751c020

Otter, D. W., Medina, J. R., & Kalita, J. K. (2021). A Survey of the Usages of Deep Learning for Natural Language Processing. *IEEE Transactions on Neural Networks and Learning Systems*, *32*(2), 604–624. https://doi.org/10.1109/TNNLS.2020.2979670

Peters, M. D. J., Godfrey, C. M., Khalil, H., McInerney, P., Parker, D., & Soares, C. B. (2015). Guidance for conducting systematic scoping reviews. *International Journal of Evidence-Based Healthcare*, *13*(3), 141–146. https://doi.org/10.1097/XEB.0000000000000050

Peters, M. D. J., Godfrey, C., Mcinerney, P., Baldini Soares, C., Khalil, H., & Parker, D. (2017). Chapter 11: Scoping reviews. In E. Aromataris & Z. Munn (Eds.), *Joana Briggs Institute Reviewer's Manual* (Issue September, pp. 6–24). https://reviewersmanual.joannabriggs.org/

Pictet, J. (1979). *The fundamental principles of the Red Cross*. http://www.ifrc.org/PageFiles/40669/Pictet Commentary.pdf

Ramalingam, B., Scriven, K., & Foley, C. (2009). *Innovations in international humanitarian action: ALNAP's 8th review of humanitarian action*. ALNAP. https://www.alnap.org/help-library/innovations-in-international-humanitarian-action-alnaps-8th-review-of-humanitarian

Raymond, N. A., Al Achkar, Z., Verhulst, S., & Berens, J. (2016). Building data responsibility into humanitarian action. *OCHA Policy and Studies Series*, *018*, 18. https://docs.unocha.org/sites/dms/Documents/TB18_Data Responsibility_Online.pdf

Redmond, A. D. (2005). Needs assessment of humanitarian crises. *Bmj*, *330*(7503), 1320–1322.

https://doi.org/10.1136/bmj.330.7503.1320

Roberts, T., & Faith, B. (2021). *Digital Aid: understanding the digital challenges facing humanitarian assistance*. *March*. https://doi.org/10.19088/IDS.2021.030

Sarı, L., Hasegawa-Johnson, M., & Yoo, C. D. (2021). Counterfactually Fair Automatic Speech Recognition. *IEEE/ACM Transactions on Audio Speech and Language Processing*, *29*, 3515–3525. https://doi.org/10.1109/TASLP.2021.3126949

Savoldelli, A., Codagnone, C., & Misuraca, G. (2014). Understanding the e-government paradox: Learning from literature and practice on barriers to adoption. *Government Information Quarterly*, *31*(SUPPL.1), S63–S71. https://doi.org/10.1016/j.giq.2014.01.008

Sheather, J., Jobanputra, K., Schopper, D., Pringle, J., Venis, S., Wong, S., & Vincent-Smith, R. (2016). A Médecins Sans Frontières Ethics Framework for Humanitarian Innovation. *PLoS Medicine*, *13*(9), 1–10. https://doi.org/10.1371/journal.pmed.1002111

Shoemaker, E., Kristinsdottir, G. S., Ahuja, T., Baslan, D., Pon, B., Currion, P., Gumisizira, P., & Dell, N. (2019). Identity at the margins: Examining refugee experiences with digital identity systems in Lebanon, Jordan, and Uganda. *COMPASS 2019 - Proceedings of the 2019 Conference on Computing and Sustainable Societies*, 206–217. https://doi.org/10.1145/3314344.3332486

Slim, H. (2015). *Humanitarian Ethics: A Guide to the Morality of Aid in War and Disaster*. Oxford University Press.

Smith, A., Pringle, J., & Hunt, M. (2020). Value-Sensitive Design for Humanitarian Action: Integrating Ethical Analysis for Information and Communication Technology Innovations. In D. Messelken & D. Winkler (Eds.), *Ethics of Medical Innovation, Experimentation, and*

Enhancement in Military and Humanitarian Contexts (pp. 105–124). Springer. https://doi.org/10.1007/978-3-030-36319-2_7

Spiegel, P. B., Burkle, F. M., Dey, C. C., & Salama, P. (2001). Developing public health indicators in complex emergency response. *Prehospital and Disaster Medicine*, *16*(4), 281–285. https://doi.org/10.1017/S1049023X00043430

Tatman, R. (2017). Gender and Dialect Bias in YouTube's Automatic Captions. *Proc. of the First Workshop on Ethics in Natural Language Processing*, 53–59. http://www.aclweb.org/anthology/W17-1606

Tatoeba. (2022). *Tatoeba*.

The Engine Room and Oxfam. (2018). *Biometrics in the Humanitarian Sector*. https://www.theengineroom.org/wp-content/uploads/2018/03/Engine-Room-Oxfam-Biometrics-Review.pdf

Travis, D., & Hodgson, P. (2019). *Think Like a UX Researcher: How to Observe Users, Influence Design, and Shape Business Strategy* (1st ed.). CRC Press.

UN Global Pulse. (2016). *Big Data for Development and Humanitarian Action: Towards Responsible Governance: Global Pulse Privacy Advisory Group Meetings*. http://unglobalpulse.org/sites/default/files/Big_Data_for_Development_and_Humanitarian_Action_Report_Final_0.pdf

UNHCR. (2006). *Master Glossary of Terms Rev. 1 United Nations High Commissioner for Refugees*. Geneva: UN High Commissioner for Refugees (UNHCR). http://www.refworld.org/docid/42ce7d444.html

UNHCR. (2017). *Needs Assessment Handbook*.

http://needsassessment.unhcr.org/media/downloads/2017-05_NAH_BOOK_NEW.pdf

Vaishnavi, V. K., & Kuechler, W. (2015). Chapter 2: Introduction to Design Science Research in Information and Communication Technology. In *Design science research methods and patterns: Innovating information and communication technology* (2nd ed.). CRC Press. https://doi.org/10.1201/9781420059335

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). *Attention Is All You Need*. https://arxiv.org/pdf/1706.03762.pdf

Weiss, R. J., Chorowski, J., Jaitly, N., Wu, Y., & Chen, Z. (2017). Sequence-to-sequence models can directly translate foreign speech. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH, 2017-Augus*, 2625–2629. https://doi.org/10.21437/Interspeech.2017-503

WHO. (2007). *Risk reduction and emergency preparedness*. World Health Organization. https://apps.who.int/iris/bitstream/handle/10665/43736/9789241595896_eng.pdf

XLSForm.org. (n.d.). *XLSForm.org*.

6.6. Chapter 6 - Appendix 1

Tasks and Prompts for Each Round of User Testing

6.6.1. User Testing Round 1

Instructions

- Let's assume you have been put in charge of reviewing this KoboToolbox survey. The survey includes a number of audio files, and you were asked to transcribe these files as well as translate them if necessary. Let's assume all of this can be done within KoboToolbox, without having to download the audio files first.

Activity 1 - Finding functionality

- Where would you find these audio files?
- Please start with [fictional respondent name]. We want to start with him. How would you find, open, and start transcribing his audio?
- Now let's imagine you are already done with [fictional respondent name] and you want to go to the audio of [other fictional respondent name], how would you get there?

Activity 2 - Transcribing a submission

- Now that you are in the processing view for [fictional respondent name]'s audio file, you want to transcribe his audio. It's in Italian. You want to do it automatically.
- Great! Now let's say that your rough Italian is enough to spot a few errors here and there. How would you edit the text?
- What if you think it's a terrible transcription and you wanted to delete it and start over again?

Activity 3 - Translating submission

- Now that you have your transcript in Italian, you want to make sure the team at HR can read this testimony, so you need to translate it to English. Since your Italian is not so good, you want to do this using the automated option again, but you want to make sure it's translated to English from the UK, since that's where HQ of the org are.
- Now being an international NGO, you also want to translate this to French. Fortunately, you speak both English and French very well, so you want to do this translation, from English to French, manually. How would you do this?

Feedback

- What did you think of the functionality? Did you find it intuitive?
- What aspects of the activities did you struggle with? Is there anything you would have expected to be different?
- Would this be useful for your work? Would you use it?
- Do you have any other feedback or suggestions?

6.6.2. User Testing Round 2

Instructions

Let's assume you have been put in charge of an internal survey at the Global Education Cluster. The goal of the survey is to better understand how each of our team members feel about the place they live and work from. Your supervisor asked you to make sure that all the audio questions are transcribed correctly, and to translate them into English if necessary. They told you to do all that in KoboToolbox.

Activity 1 - Finding audio

- Where would you find these audio files?

- Please start with [AUDIO QUESTION 1] from the submission from [NAME] What language do you think this audio is in?

  [Use submission relevant to speaker of that language]

- How would you transcribe this audio file?

Activity 2 - Transcription EN

- Please go ahead with the **transcription** of the audio file.
- (If participant chooses ASR): How would you make sure that the **transcript is accurate**?

Activity 3 - Translation EN to target language (TG)

- How would you go about **translating** this transcript into (TG)?
- (If participant chooses MT): How would you make sure that the **translation is accurate**?

Activity 4 - Transcription TG

- Now that you have transcribed and translated the first response, please **transcribe** the second question for the next submission.
- [If participant chose ASR in Activity 2, ask them to do manual transcription, or vice versa]

Activity 5 - Translation TG to EN

- Please go ahead and **translate** the second transcript into English.
- [If participant chose MT in Activity 3, ask them to do manual translation, or vice versa]
- Save your work and go back to the table

Activity 6 - Review and download data

- How would you **review** the different transcripts you've created so far?
- How would you **download** the different transcripts?

Feedback

- How useful do you think this functionality would be for your work? How would you use it exactly?
- What aspects of the activities did you struggle with? Is there anything you would have expected to be different?
- Is there anything you found was missing?

Questions about qualitative analysis

- If you were to collect this type of data for a large-scale project, how would you analyze/categorize it once you have these transcripts/translations?
- Are you already using other tools for this? Which ones?
- What would be the ideal outcome of these transcripts? What would you hope to get from it?
- What would you like to see in this interface that could help you do that (analyze)?
- Would you like to see some of these features here, or would you rather do this type of work in other tools?

6.7. Chapter 6 - Appendix 2

Proposed Conceptual Design of NLP Features

6.7.1. Personas

Marcela

- Assessment Coordinator at Swiss Refugee Committee (SRC) for 10 years
- Worked in 20 different humanitarian crises
- Wants a standard approach for qualitative data to help measure needs
- Finds qualitative interview notes hard to read and unreliable

Gabriele

- Translator working in Guatemala City
- Works for SRC since 2011
- Fluent in Spanish and English
- Has used some tools for transcribing in English but mostly does it manually

Linh

- M&E specialist, consultant for SRC
- Works with aid organizations to conduct large assessment and research projects in crisis environments
- Records interview responses as audio and quantitative data with KoboToolbox but never together

6.7.2. User Stories

Marcela

- As survey administrator, I want to make sure that written transcriptions and translations from audio responses accurately capture the basic meaning so that our intervention provides the right kind of aid.
- As Emergency Assessment Coordinator, I want all field teams to move away from using handwritten notes for qualitative questions and use full transcripts/translations instead.

Gabriele

- As a translator, I want to quickly create transcripts and translations from audio data so that Marcela gets the information she wants and people in my country get the help they need.

Linh

- As an M&E specialist, I want to collect qualitative information to understand the opinions of refugees of SCR's work in the country so that I can create a detailed report about issues that should be improved.
- Like Marcela, I want to have written transcripts and translations from audio responses

6.7.3. Scenario

A conflict has erupted on the border with Country X. Refugees are arriving in great numbers.

Marcela is sent to the country to lead an existing SRC team to do a rapid needs assessment. After three days on the ground, she launches a survey with KoboToolbox containing closed-ended and open-ended qualitative questions. The qualitative questions require enumerators to record responses as audio within KoboCollect.

6.7.4. Proposed User Flow

Create form

7. Marcela creates form in KoboToolbox using a combination of text, select_one, integer, as well as 4 audio questions. The survey will be conducted in Spanish.
8. Data is being collected by a team of 10 enumerators over 5 days.

Collect audio data in the field

- User collects non-audio questions
- Audio question is displayed in KoboCollect
- User selects a response while recording the audio (to get an immediate response coding)
- User moves on to other questions

Automated Transcription

1. Marcela logs into her account to view data collected so far.
2. She opens the Table view and sees collected data. For audio questions she sees a ▶ button in each row which she uses to play back the respective audio file.

3. She wants to enable automatic transcription from audio to Spanish text for each of the audio questions:
    a. She clicks on “Transcription/translation” in the left menu (or clicks on the ‘Transcribe/Translate’ button above one of the audio questions)
    b. A new screen opens. She sees a list of her 4 audio questions.
    c. She selects all 4 audio questions
    d. In a dropdown, she selects “Spanish (Guatemala)” as the source language.
    e. In a dropdown, she keeps the default transcription method as “Google Cloud”
    f. She ticks “Enable automatic transcription for future audio as well”
    g. She clicks on “Enable transcription”
    h. A message tells her: “Transcription started. Based on the number of files and length this may take a while . Please check back in a few minutes to see transcriptions arrive in the Table view. Based on your setting above, all future audio will be transcribed once received by the server.”

Automated Translation

4. Marcela now wants to enable automatic translation
    a. She clicks on “Transcription/translation” menu item
    b. She again selects her 4 audio questions
    c. In a dropdown, she selects “English (US)” as the target language.

   d. In a dropdown, she changes the default translation method to “Amazon Web Services”
   e. She keeps the check on the default option “☑ Only translate transcriptions that have been marked as verified”
   f. She clicks “Enable translation”
5. After half an hour she opens the Table view and sees that her audio questions now have text next to the ▶ button.
6. She shares her project with Gabriela (giving her Edit rights) so that she can correct the automated transcriptions. She calls her to let her know to start working on this.

Corrections to Automated Transcription or Translation

1. Gabriele signs in to her account to correct the transcriptions that Marcela asked her to work on.
2. She opens the project, clicks Data, clicks on “Transcription/Translation”
3. She clicks on “View and correct transcriptions and translations”
4. Gabriele selects the first of the 4 audio questions (“¿Cómo su familia encuentra comida desde que dejó su hogar?”)
5. She sees a table with two columns (transcription, translation)
   a. In each row under the transcription field there is a ▶ button and the transcription text.

b. She clicks the ▶ button and listens to the recording, reading along the text that was transcribed automatically.

c. Underneath the text field, a slider and a time indicator shows progress (e.g. 0:09 / 3:01)

d. Gabriele clicks the ❚❚ button to pause and catch up with her edits, then ▶ again until she is done with the particular response.

e. She now marks the transcription as “verified”.

f. An automated translation text appears in the second column.

g. She verifies the translation, makes edits, and then marks the translation as “verified”.

h. She then goes to the next row, repeating the above steps until the first audio question is completed.

i. She then moves on to the other 3 questions.

6. When Gabriele is done, the “View and correct transcriptions/translations” screen shows each of the 4 questions as “100% verified”
7. She sends an email to Marcela to tell her she is done.

Verify and Download Data

1. Marcela opens the Table view of her project to create translations of the transcriptions.
2. She sees new columns next to each audio question for the transcriptions and translations.

3. She downloads the data as XLS and begins her analysis in Excel.

Chapter 7 Conclusion

The purpose of this dissertation research was to investigate whether the design of new software features for KoboToolbox, a computer-assisted personal interviewing (CAPI) tool, can result in better data for humanitarian needs assessments (HNAs), that more accurately reflect the needs of people affected by humanitarian crises. These new CAPI software features were designed to use natural language processing (NLP), a type of artificial intelligence (AI) to systematically transcribe and translate responses to open-ended questions (OEQs). The research also aimed to understand the potential ethical implications related to processing data from affected people with a view to mitigate against ethical issues in the new CAPI features. This manuscript-based dissertation used a mixed-method Design Science Research approach rooted in the pragmatist philosophy. It did so by conducting key informant (KI) interviews, a scoping review of the literature, designing new CAPI features based on a user-centered design method, and by collecting feedback from test users about the new features' usability.

Chapter 1 described the rationale for this dissertation research, laid out key definitions, an overview of the literature, and the research questions, and described the York University's ethical review and approval of the study.

Chapter 2 described the theoretical frameworks and provided an overview of the methodologies applied. It explained the use of the Design Science Research approach and showed how this dissertation is grounded in the pragmatist philosophy, as well as in a set of values in the humanitarian assistance, bioethics, and design science sectors, respectively. The findings of this dissertation research were laid out in four interconnected studies in chapters 3, 4,

5, and 6, using a format that is intended for publication in peer-reviewed academic journals, as explained below.

Chapter 3 presented a chapter manuscript entitled “Improving Humanitarian Needs Assessments through Natural Language Processing.” It described how NLP could be used to transcribe, translate, and analyze large sets of qualitative responses with a view to improving the quality and effectiveness of humanitarian assistance, while also outlining the anticipated practical and ethical challenges of such an innovation.

Chapter 4 introduced a chapter manuscript titled *Overcoming Blind Spots: Constraints and Solutions Related to Qualitative Interview Methods in Humanitarian Needs Assessments*. This study explored current approaches to conducting HNA, constraints related to different interview methods—particularly qualitative ones—and solutions for improving the use of qualitative interview methods in HNAs.

The chapter manuscript included as Chapter 5 is entitled *Ethical Implications Related to Processing of Personal Data in Humanitarian Crises: A Scoping Review*. This study mapped the range of ethical issues that have been raised in the academic literature regarding the processing of data from people affected by humanitarian crises.

The Chapter 6 manuscript is titled *Systematic Design and Evaluation of New Humanitarian Needs Assessments Tools for Collecting Qualitative Data Using Natural Language Processing*. This study reports on the design and empirical evaluation of new features added to KoboToolbox, a CAPI tool, to systematically transcribe, translate, and analyze answers to OEQs gathered from HNA interviews.

This Conclusion (Chapter 7) ends the dissertation with a review of the methodologies and an overview of the key findings to each of the research questions. It then discusses the

significance of the study for policy and practice, and its original contribution to the academic literature. Finally, this chapter summarizes this study's methodological limitations and provides recommendations for future research, before offering a final reflection on the overall research process.

7.1. Key Findings

This section summarizes the research questions and chosen methodologies before providing an overview of the key findings in relation to each of the three areas of study described in the literature review in Chapter 1.

7.1.1. HNA Practices and Methodological Constraints

In order to investigate the prevailing practices and methodological challenges in HNA, the following research questions were identified:

4. What are the current approaches to conducting HNAs in a sample of professional staff members of humanitarian organizations with prior experience managing HNA data collection?
5. Among these respondents, what constraints related to different interview methods, particularly qualitative ones, are identified?
6. Which solutions for improving the use of qualitative interview methods in HNAs do respondents propose?

As detailed in Chapter 4, this part of the dissertation was based on semi-structured interviews with 23 KIs who had prior experience collecting data for HNA. KIs were drawn form from 13 countries and represented 17 different humanitarian organizations, including United Nations (UN) agencies, the Red Cross/Red Crescent Movement, as well as nongovernmental

organizations (NGOs). Analysis of the transcribed interviews resulted in 47 themes, which were further organized into five Thematic Groups: 1) Prevailing approaches to data collection in HNA; 2) constraints related to quantitative interview methods; 3) constraints related to qualitative interview methods; 4) challenges specific to pre-coded open-ended questions; and 5) respondents' recommendations for improving the role of qualitative interview methods in HNA.

The findings showed that current approaches to conducting HNA often favor quantitative methods, particularly household surveys for multi-sector needs assessments (MSNA). Even where qualitative data are collected, they are often not fully analyzed or included in HNA reporting. KIs pointed to numerous constraints to collecting qualitative data in HNA due to resource shortages, insufficient training of relevant staff members, lack of time to analyze qualitative data, and engrained biases in favor of quantitative data. As a result, the role of qualitative methods in HNA remains circumscribed compared to that of quantitative methods—even though qualitative data were widely seen by KIs as crucial for understanding population needs. Significant challenges were identified to using pre-coded response scales in CAPI survey instruments as a way to capture the information elicited from OEQ. Proposals by KIs to close this gap included: 1) acknowledging the limitations of quantitative data to informing HNA; 2) investing in more qualitative skills through training or hiring; 3) piloting mixed-methods hybrid surveys (a combination of closed-ended questions and OEQs for which audio would be recorded); as well as 4) employing NLP for transcribing, translating, and analyzing interview responses to OEQs.

7.1.2. Ethical Implications Related to Processing Personal Data

Regarding ethical implications related to processing data of people affected by humanitarian crises, the following research questions were posed:

7. Which ethical issues have been raised in the peer-reviewed literature related to processing data from people affected by humanitarian crises in order to inform humanitarian assistance?
8. To what extent do real-world examples of ethical issues reflect the concerns presented in the literature?
9. Which technologies were the focus of concern over these ethical issues?

As described in Chapter 5, these questions were answered through a scoping review of the academic literature to generate a broad overview of relevant evidence and for identifying gaps in the literature (Peters et al., 2015). The scoping review method used followed the approach first described by Arksey and O'Malley (2005) and further refined by Levac et al. (2010), and follows the framework maintained by the Joanna Briggs Institute (Peters et al., 2017).

After screening 8,387 papers, the scoping review yielded 100 relevant studies. An in-depth review of the included studies led to the identification of 22 ethical issues which were then grouped along the four ethical value categories of autonomy, beneficence, non-maleficence, and justice. Slightly over half of included studies (n = 52) identified ethical issues based on real-world examples. The technologies most frequently discussed in these studies included social media, crowdsourcing, and mapping tools. On the other hand, commonly used tools identified by HNA experts in Chapter 4, such as CAPI software, spreadsheets, or use of online databases were only rarely discussed—or not at all. Various actual or potential uses of artificial intelligence were discussed in 25 studies.

Included studies showed widespread concerns that data processing in humanitarian assistance can cause additional harm, may not provide direct benefits, may limit the autonomy of

affected populations, and has the potential to lead to the unfair distribution of resources. The most-cited ethical issue (n = 74) was a concern for privacy in cases where personal or sensitive data might be inadvertently shared with third parties. Another widely discussed issue was inaccurate data that could potentially lead to a false representation of reality and therefore could misinform decisions about the provision of humanitarian assistance.

7.1.3. Design of New CAPI Features Integrating NLP

Three research questions were identified regarding the design of new CAPI features in the KoboToolbox software using NLP technologies:

10. Which specific constraints identified by professionals with experience managing HNA data collection could be addressed through new CAPI features?
11. Which ethical issues identified in the peer-reviewed literature related to processing data from people affected by humanitarian crises should be considered when designing new CAPI features for processing qualitative data?
12. How should new features for the KoboToolbox CAPI software to systematically transcribe, translate, and analyze answers to OEQs from HNA interviews be designed, usability tested, and implemented?

As described in Chapter 6, these research questions were addressed through a three-stage, mixed-methods approach that included data from qualitative interviews with KIs and the scoping review of ethical issues (presented in more detail in Chapters 4 and 5, respectively), as well as an iterative design process that was informed by and validated with usability testing. Two rounds of usability testing were conducted with a total of 14 test users residing in 13 countries.

Specific analysis of the data from interviews with KIs led to the identification of 25 themes about constraints or proposals that could be addressed through new features in CAPI

software. Seven ethical implications were found based on an analysis of the scoping review data that could be addressed through the design of new CAPI features.

Using a user-centered design (UCD) method (Gould & Lewis, 1983; Holden & Boustani, 2021), new KoboToolbox features were designed in an iterative way to first address the themes and ethical implications identified. Feedback and observations from usability testers were then used to identify strengths and weaknesses of the new features and their design. Each round of user testing led to changes in the CAPI design proposals. The final version of the new CAPI software features was released publicly as open-source software and is now available for use by humanitarian organizations. Feedback from participants showed strong support for the applicability and usability of the newly created CAPI features, particularly due the potential of NLP methods to save time and to increase the amount of qualitative data that can be collected and analyzed in humanitarian crises.

7.2. Significance of the Study

This section discusses the potential implications of this study as they relate to: 1) the practice of HNA; 2) the mitigation against ethical issues when processing data of people affected by humanitarian crises; and 3) the design of new software and other innovations for use in humanitarian assistance.

7.2.1. HNA Practices and Methodological Constraints

Based on interviews with KIs, four specific recommendations were identified that could have a wide-ranging impact on the practice of conducting HNA. These proposals were:

1. Acknowledge limitations of quantitative data
2. Increase number of staff with qualitative skills

3. Methodological innovation through hybrid surveys
4. Technological innovations

These proposals, discussed in more detail in Chapter 4, can be implemented independently of each other. The new features designed and implemented as part of this study (as described in Chapter 6) are expected to have a significant impact in addressing proposal 4. Adoption of the new features by humanitarian organizations can be expected to be rapid given the strong interest expressed by KIs, user testers, as well as by practitioners who laid out the initial need to identify better solutions for processing qualitative data in HNA (as detailed in Chapter 1). Given the widespread use of KoboToolbox by humanitarian organizations to date and the user-friendly design of the new features, it is realistic to expect few technological obstacles to the collection of more qualitative data in HNA.

As such, the most important work remaining to address proposals (1) and (2) will require institutional policy changes. Implementing these two proposals will require a concerted effort both by humanitarian organizations and by governments and international organizations who finance and control a large proportion of humanitarian assistance today (Global Humanitarian Assistance, 2021).

Proposal (3) for conducting hybrid surveys may be helpful not only for the conduct of HNA but also to researchers and practitioners involved in primary data collection for other purposes in humanitarian assistance and related sectors, such as public health or economic development. In particular, this proposal includes four steps:

1. Household or KI interview instruments should be designed to include more OEQs while reducing the number of closed-ended questions.

2. Audio of spoken responses to OEQs is recorded and quantitative data to closed-ended questions are stored together in a CAPI tool supporting such a method.
3. Rapid transcripts should be created as quickly as possible by native speakers and supported by NLP and saved as part of the survey dataset.
4. Transcripts are coded and analyzed rapidly by staff with the resulting data becoming part of the survey dataset. If possible, local staff should do this work in the interview language to avoid challenges with translating original responses.

Once implemented, it is expected that the proposals identified in Chapter 4 would improve the quality of HNAs and therefore have the potential of making humanitarian assistance more relevant, timely, and adapted to local requirements by increasing the amount of qualitative data that can be collected and analyzed more quickly and from a larger set of relevant informants.

7.2.2. Ethical Implications of Data Processing

The ethical implications identified in this study, as detailed in Chapter 5, have the potential to inform the development of ethical codes of conduct for both designing and using technology to process data from people affected by humanitarian crises. Although several such guidelines exist already, as explained in Chapters 3 and 5, to this none have been informed by the full range of ethical implications documented in the literature. It can be hoped therefore that future versions of existing guidelines will consider using the ethical implications from this study as a way to inform mitigation steps. The methods and results from this study are also expected to be of important value for future scoping reviews (or a similarly systematic approach) to identify emerging ethical implications.

The list and structuring of ethical implications identified in this study are also expected to be useful to creating an approval process. Such an approach, possibly through a checklist, could be created to require anyone involved in creating innovations for the humanitarian sector to proactively assess how a solution is designed to avoid or mitigate potential ethical implications. This could be achieved, for example, by building on the Ethics for Humanitarian Innovation Toolkit published by Elrha and Humanitarian Health Ethics (2021; Krishnaraj et al., 2021).

Finally, as proposed in Chapter 5, there is a need for more training and appropriate accountability mechanisms to monitor the actual harm or potential for causing harm. It can be expected that humanitarian organizations will be better equipped to conduct such work by utilizing the list of ethical implications identified in this study.

7.2.3. Design of New CAPI Features Integrating NLP

Three sets of important contributions from this research can be identified, and relate to the design of new CAPI features integrating NLP techniques.

First, the design and implementation of new features for the KoboToolbox CAPI software have resulted in a practical tool that addresses functional requirements identified by humanitarian professionals. This tool can now be readily deployed by humanitarian organizations to systematically transcribe, translate, and analyze answers to OEQs from HNA interviews.

Second, the research and design activities conducted for this study are expected to result in additional software features for improving the use of qualitative data in the coming years. For example, further work is expected to take place in 2023 to implement manual and automated analysis features in KoboToolbox based on transcripts created from audio files. Future software

development is also planned to enable users to access automated transcription for languages that are not yet available commercially through an integration with open-source automated speech recognition (ASR) models, as well as ASR and machine translation (MT) from other commercial providers such as IBM, Amazon Web Services, and Microsoft Azure.

Third, the multi-stage design science approach rooted in pragmatism as employed by this study has the potential to guide other software innovations for the humanitarian assistance sector. Humanitarian organizations adopting the UCD method used in this study may be better equipped to address ethical implications and prevent abrupt institutional measures such as Oxfam's 2015 decision to temporarily stop the use of all biometric systems, due to unaddressed ethical concerns (Eaton-Lee & Shaughnessy, 2021). This study showed that it is possible to inform designs at the earliest stage by a review of potential ethical issues. Adopting such a method of putting ethical considerations at the beginning of the design cycle can have a significant impact, therefore, to guard against future ethical risks, particularly in the humanitarian and health sectors.

By documenting the UCD process of creating significant new features for a widely used software in the humanitarian assistance sector, this study also hopes to serve as a strong encouragement to adopt such a process more widely in practice. Doing so may help prevent severe usability issues that can lead to low rates of implementation and therefore reduce the high costs of implementing significant changes later on.

As demonstrated in the literature (Cornet & Holden, 2018; Fiordelli et al., 2013), there is often a misalignment between the rigorous boundaries imposed by risk adverse internal cultures, and the flexibility required during the design and creation stages that is central to creating effective software. This phenomenon may be particularly acute for some humanitarian organizations due to administrative requirements imposed by donor funding, or due to internal

procurement rules. It is well recognized in the software development sector that those rules require new innovations to be well defined in scope, budget, and a deliverable timeframe—often years before a final tool can be launched. Such rigid structures may in turn incentivize an overly bureaucratic and risk adverse method of creating innovations, rather than encouraging a flexible design process whose goal is to respond meaningfully to user needs. As a result of an inflexible approach, humanitarian organizations may not stand to benefit from the rapid improvements found in many platforms created by private sector companies that rely on UCD and agile innovation methods to create tools that best respond to user needs (Holden et al., 2021). Rigidity can also endanger the long-term sustainability of tools created in such a fashion: Many humanitarian innovations face a lack of funding to cover the high ongoing costs for maintaining them once the initial pilot was realized (Currion, 2020; Finnigan & Farkas, 2020).

In recent years, donor initiatives such as *Humanitarian Grand Challenges* (Humanitarian Grand Challenges, 2022), the *Humanitarian Innovation Fund* (Elrha, 2022), or the World Food Programme's *Innovation Accelerator* (WFP, 2022) have sought to introduce less rigid approaches to change and innovation by encouraging grantees to use to UCD for designing their innovations. However, based on observations by the researcher, there has been only limited success to create the environment and administrative rules required to foster the use of UCD more widely.

This study hopes to contribute to an acceleration of UCD in the humanitarian sector. Chapter 6 provides a detailed example of how initial designs were changed iteratively, and final outputs were informed through user testing rather than through up-front specifications documented in contractual terms. It is hoped that more humanitarian organizations—as well as donor governments and grant-making institutions—will better recognize the value of conducting

technological innovation in such an iterative fashion. Putting the approach of this study into practice will mean accepting that assumptions about user needs or the initial design may have flaws that lead to the development of a tool that is different from what was initially envisaged. It may also require flexibility in addressing at least one dimension (funding, timing, or scope) as usability testing identifies previously undiscovered issues that can significantly alter original plans (Cornet et al., 2020).

7.3. Original Contributions of the Study

This section considers the original contributions of this study to three different areas of scholarship. They include: 1) establishing an in-depth analysis about the practice and methodological approaches of HNA; 2) creating a systematic review of ethical implications related to processing data of affected people; 3) introducing a Design Science Research approach using UCD to the field of humanitarian assistance; and 4) documenting an applied method for integrating ASR and MT capabilities within widely used CAPI software.

First, the results from this study's KI interviews fill an important gap in the academic literature. Based on a review of the available evidence, this is the first academic study that provides empirical data about practices and challenges in HNA. The strengths of this research include a diverse sample of KIs with deep experience in HNA from an array of organizations and a large geographic area, as well as a detailed investigation of several connected topics that ranged from institutional constraints to the technical abilities of different software. In particular, the difficulties documented in this study about the collection, analysis, and reporting of qualitative data have not previously been published. The findings from Chapter 4 also mark a significant contribution to the evidence on the challenges inherent in HNA. This gap in knowledge was previously identified as hindrance to implementing the commitments articulated

in the Grand Bargain to improve needs assessments and to include affected people in decision-making (WHS, 2016).

Second, despite the significant amount of interest in, and concern about, ethical implications in the literature, this study appears to be the first comprehensive review of relevant ethical implications about processing data from affected people for the purpose of humanitarian assistance. By choosing to search a wide array of databases and using a broad set of keywords, this study appears to have overcome some of the limitations of other scoping reviews. This study therefore connects a growing multidisciplinary area of research that has suffered from a lack of common terminologies and shared academic approaches. This was manifested in the significant challenges in trying to identify the relevant published literature by searching for terms that could entail the three concepts of "humanitarian assistance," "processing data," and "ethical implications." This study's search strategy made every effort to represent the most comprehensive and inclusive set of keywords, and to capture studies in each of the three identified areas of scholarship to date. Further, by documenting detailed technical decisions on whether certain events constituted a "humanitarian crisis," this study provides a transparent method for including or excluding different countries or contexts, rather than leaving this up to personal interpretations during the screening process.

Third, the results from this study's design and empirical user testing fill an important gap in academic research about the practical design of software and other innovations for humanitarian assistance. The strengths of the results documented in Chapter 6 include two rounds of user testing with a diverse and international sample of potential users, in order to identify and address usability issues. Strengths also include a systematic approach of using a

scoping review of ethical implications and interviews with HNA experts to inform the initial conceptual design.

Fourth, this study presents the first documented instance, based on a review of the available published literature, that describes the design and implementation of new CAPI software features that integrate ASR and MT capabilities. Particularly by documenting the design challenges of implementing these NLP methods in a tool that is not limited to a single purpose or language, but rather needs to function well with thousands of languages and with many types of data collection methods, this study also contributes to the related disciplines of UCD, design science, e-health, and computer science.

7.4. Limitations of the Study

Relevant limitations related to the interviews with key informants include a relative over-sampling of KIs from international humanitarian organizations, as well as under-sampling of KIs from Africa, East Asia, and Latin America. Although KIs between them represented 17 organizations, it cannot not be assumed that the findings from these interviews reflect all views within these institutions, nor that the perspectives of those KIs are representative of other humanitarian organizations.

This scoping review in this study was limited to literature published since 2010 and before January 2020, and it excludes work from non-peer-reviewed sources. As described in Chapter 5, identifying all relevant studies was a significant challenge as no shared nomenclature across disciplines for humanitarian assistance, ethical implications, and data processing could be identified. The set of keywords, created through multiple stages of piloting, was intended to be as inclusive as possible. However, potentially relevant articles that met the inclusion criteria may still have been missed.

For the design research aspect, although user testing candidates were chosen to represent the diverse community of current or future users involved in qualitative data collection, the final sample cannot be assumed to be representative of the overall population. The two user testing samples may therefore have missed divergent opinions and different feedback from other types of users or organizations.

7.5. Recommendations for Future Research

This section briefly outlines some of the areas recommended for further study that have emerged from the findings related to the fields of HNA methods, ethical implications of processing data from affected people, and designing new features for CAPI tools.

First, related to the practice of HNA, further research is needed to pilot and assess the proposed hybrid surveys method. A representative survey among HNA experts would be extremely relevant to quantify some of the themes identified in Chapter 4, particularly related to the institutional constraints that limit the collection or reporting of certain data.

The findings from this study indicate that much quantitative data is being collected without being applied to practice. Given the high costs of MSNA, which take place in virtually all protracted humanitarian crises, it is recommended that a study be conducted to assess the usefulness of collected MSNA data for different purposes and actors. Finally, most KIs indicated that a lack of enumerator training is a significant source of unreliable data that can severely impact the accuracy of HNA results. More research is needed therefore to establish minimum standards for enumerator skills for different kinds of methods—as well as a study to assess enumerator quality by rigorously accounting for the way an interview was conducted and responses were coded quantitatively. Such a study would be made significantly easier through

the use of background audio recording and ASR features that were implemented as part of this study.

Second, related to ethical implications of processing data for the purpose of humanitarian assistance, more research using empirical methods is necessary to better identify and understand the prevalence of the ethical issues identified in this study. Further, as the findings in Chapter 5 show, not enough is known about the heightened ethical risks stemming from data processing that may affect vulnerable people in the large number of humanitarian crises associated with conflict, war, and social instability. More research is therefore urgently needed to investigate ethical issues that arise specifically in conflict settings.

A representative study would be extremely useful to quantify the awareness of different ethical implications among staff at humanitarian organizations who are directly responsible for processing data from affected people, and to investigate the extent to which minimum data protection standards, such as published by the International Committee of the Red Cross (2020), are applied in practice. Case studies of early adoptions of NLP and other AI methods would be useful in investigating the ethical considerations that took place when designing and deploying such tools, to discover whether new ethical implications might emerge from such early adoptions. Finally, more investigations are needed to uncover the ethical implications of using commonplace tools or methods to process data for humanitarian assistance, such as CAPI tools, spreadsheets, filesharing, or use of online databases.

Third, related to designing new software, rigorous field piloting of new KoboToolbox features for recording, transcribing, and analyzing more open-ended questions would be crucial, to evaluate the efficacy of using ASR and MT functionality. Such research would test the effectiveness of using more qualitative methods to better inform humanitarian operations and to

assess to what extent different qualitative data can impact operational decisions made by humanitarian organizations. Additional research would also be useful to evaluate the quality of transcripts created by ASR and MT based on audio recordings of survey responses created in a representative field setting.

7.6. Final Reflections

> The open ended [question] is where you get the true impact. People might speak more to their unique situation. And I think that if we had some way to sort of analyze it at that level, we would just be able to speak more to our impact. (KI #03)

This study was borne out of the cries for help by humanitarian practitioners expressed during a York University Dahdaleh Institute humanitarian conference on Humanitarian Data Science (see Chapter 1). Whereas tools such as KoboToolbox drastically changed how quantitative data could be collected to inform HNA, qualitative data from OEQs remained difficult to collect and analyze—despite the knowledge that it represented a better way understand complex realities. As a result, such data was either very limited in scope, under analyzed, or not collected in the first place. The idea of integrating NLP methods with KoboToolbox for transcribing and translating audio responses to OEQs was quickly confronted with significant design challenges as well as potential ethical risks of involving emerging AI methods, as outlined in Chapter 3.

This study's empirical chapters used a pragmatic approach to understand the methodological practices of HNA experts, systematically assess relevant ethical issues, as well as design and empirically evaluate new CAPI features for collecting and transcribing audio

responses to OEQs. It is hoped that the newly designed features described in Chapter 6 will have a significant impact in the conduct of HNA and other types of evidence collection in public health and related sectors.

These features can only address but a small set of the constraints identified in Chapter 4 and play a modest role in mitigating potential ethical risks documented in Chapter 5. Yet, in conclusion, this study hopes to also impact future innovations created to improve humanitarian assistance by demonstrating that a pragmatic iterative design process is not only feasible in the humanitarian sector, but also likely generates better results. As the drivers behind the significant increase in humanitarian needs over the last years are expected to continue or even accelerate, global resources dedicated to providing humanitarian assistance will likely continue to be strained. In this context, the innovations focused on HNAs that are described in this dissertation may make the provision of life saving humanitarian assistance more effective in the future.

7.7. References

Arksey, H., & O'Malley, L. (2005). Scoping studies: Towards a methodological framework. *International Journal of Social Research Methodology*, *8*(1), 19–32. https://doi.org/10.1080/1364557032000119616

Cornet, V. P., & Holden, R. J. (2018). Systematic review of smartphone-based passive sensing for health and wellbeing. *Journal of Biomedical Informatics*, *77*(October 2017), 120–132. https://doi.org/10.1016/j.jbi.2017.12.008

Cornet, V. P., Toscos, T., Bolchini, D., Ghahari, R. R., Ahmed, R., Daley, C., Mirro, M. J., & Holden, R. J. (2020). Untold stories in user-centered design of mobile health: Practical challenges and strategies learned from the design and evaluation of an app for older adults with heart failure. *JMIR MHealth and UHealth*, *8*(7), 1–14. https://doi.org/10.2196/17703

Currion, P. (2020). The Black Hole of Humanitarian Innovation. *Journal of Humanitarian Affairs*, *1*(3), 42–45. https://doi.org/10.7227/jha.024

Eaton-Lee, J., & Shaughnessy, E. (2021, June 25). *Oxfam's new policy on biometrics explores safe and responsible data practice*. Reliefweb. https://reliefweb.int/report/world/oxfam-s-new-policy-biometrics-explores-safe-and-responsible-data-practice

Elhra, & Humanitarian Health Ethics. (2021). *Ethics for humanitarian innovation: Toolkit*. https://higuide.elrha.org/wp-content/uploads/2021/09/Elrha-Ethics_for_Humanitarian_Innovation-Toolkit-210913.pdf

Elrha. (2022). *Humanitarian Innovation Fund*. https://www.elrha.org/programme/hif/

Finnigan, G., & Farkas, O. (2020). More than Laboratories: Four Decisive Challenges Confronting Humanitarian Innovation. *Journal of Humanitarian Affairs*, *1*(3), 4–13.

https://doi.org/10.7227/jha.020

Fiordelli, M., Diviani, N., & Schulz, P. J. (2013). Mapping mhealth research: A decade of evolution. *Journal of Medical Internet Research*, *15*(5). https://doi.org/10.2196/jmir.2430

Global Humanitarian Assistance. (2021). *Global Humanitarian Assistance Report 2021*. Development Initiatives. https://devinit.org/documents/1008/Global-Humanitarian-Assistance-Report-2021.pdf

Gould, J. D., & Lewis, C. (1983). Designing for usability-key principles and what designers think. *Conference on Human Factors in Computing Systems - Proceedings*, *28*(3), 50–53. https://doi.org/10.1145/800045.801579

Holden, R. J., & Boustani, M. A. (2021). Design and Usability Methods. In R. J. Holden & R. S. Valdez (Eds.), *The Patient Factor* (1st ed.). CRC Press.

Holden, R. J., Boustani, M. A., & Azar, J. (2021). Agile Innovation to transform healthcare: Innovating in complex adaptive systems is an everyday process, not a light bulb event. *BMJ Innovations*, *7*(2), 499–505. https://doi.org/10.1136/bmjinnov-2020-000574

Humanitarian Grand Challenges. (2022). *Creating Hope in Conflict: A Humanitarian Grand Challenge*. https://humanitariangrandchallenge.org/

International Committee of the Red Cross. (2020). *Handbook on Data Protection in Humanitarian Action (2nd Edition)*. https://www.icrc.org/en/data-protection-humanitarian-action-handbook

Krishnaraj, G., Hunt, M., Schwartz, L., O'Mathúna, D., Pringle, J., Okhowat, A., Kapiriri, L., Brahimi, L., & Yantzi, R. (2021). *Ethics for Humanitarian Innovation: Background Paper*. Elrha. https://higuide.elrha.org/wp-content/uploads/2021/09/Elrha-

Ethics_for_Humanitarian_Innovation-Background_Paper-210909.pdf

Levac, D., Colquhoun, H., & O'Brien, K. K. (2010). Scoping studies: Advancing the methodology. *Implementation Science*, *5*(1). https://doi.org/10.1186/1748-5908-5-69

Peters, M. D. J., Godfrey, C. M., Khalil, H., McInerney, P., Parker, D., & Soares, C. B. (2015). Guidance for conducting systematic scoping reviews. *International Journal of Evidence-Based Healthcare*, *13*(3), 141–146. https://doi.org/10.1097/XEB.0000000000000050

Peters, M. D. J., Godfrey, C., Mcinerney, P., Baldini Soares, C., Khalil, H., & Parker, D. (2017). Chapter 11: Scoping reviews. In E. Aromataris & Z. Munn (Eds.), *Joana Briggs Institute Reviewer's Manual* (Issue September, pp. 6–24). https://reviewersmanual.joannabriggs.org/

WFP. (2022). *Innovation Accelerator*. https://innovation.wfp.org/

WHS. (2016). *The Grand Bargain – A Shared Commitment to Better Serve People in Need*. Istanbul: World Humanitarian Summit. http://reliefweb.int/sites/reliefweb.int/files/resources/Grand_Bargain_final_22_May_FINAL-2.pdf

Bibliography

Abbott, J. Z., & Martinus, L. (2018). Towards Neural Machine Translation for African Languages. *ArXiv:1811.05467 [Cs, Stat]*. https://arxiv.org/abs/1811.05467v1

ACAPS. (2011). *Technical brief: Direct observation and key informant interview techniques for primary data collection during rapid assessments.* ACAPS. https://www.acaps.org/sites/acaps/files/resources/files/direct_observation_and_key_informant_interview_techniques_for_primary_data_collection_during_rapid_assessments_october_2011.pdf

ACAPS. (2012). *Qualitative and quantitative research techniques for humanitarian needs assessment*. ACAPS. https://www.acaps.org/sites/acaps/files/resources/files/qualitative_and_quantitative_research_techniques_for_humanitarian_needs_assessment-an_introductory_brief_may_2012.pdf

ACAPS. (2013). *How sure are you ?* ACAPS. https://www.acaps.org/sites/acaps/files/resources/files/how_sure_are_you-judging_quality_and_usability_of_data_collected_during_rapid_needs_assessments_august_2013.pdf

ACAPS. (2014). Humanitarian Needs Assessment: The Good Enough Guide. In *Geneva: ACAPS*. ACAPS. https://doi.org/10.3362/9781780449357

ACAPS. (2016). *Meeting information needs? A review of ten years of multisector coordinated needs assessment reports*. Geneva: ACAPS. https://www.acaps.org/sites/acaps/files/resources/files/acaps_report-a_review_of_ten_years_of_multisector_coordinated_needs_assessment_reports_january_2016.pdf

ACAPS. (2019). *Lessons Learned: Needs Assessments in Cox's Bazar*. ACAPS. https://www.acaps.org/sites/acaps/files/products/files/20190404_acaps_npm_coxs_bazar_analysis_hub_needs_assessments_lessons_learned.pdf

Adikari, S., McDonald, C., & Campbell, J. (2009). Little design up-front: A design science approach to integrating usability into agile requirements engineering. *Human-Computer Interaction. New Trends. HCI 2009. Lecture Notes in Computer Science*, *5610*, 549–558. https://doi.org/10.1007/978-3-642-02574-7_62

AI Now Institute. (2018). *AI in 2018: A year in review*. Medium. https://medium.com/@AINowInstitute/ai-in-2018-a-year-in-review-8b161ead2b4e

Al-Amer, R., Ramjan, L., Glew, P., Darwish, M., & Salamonson, Y. (2016). Language translation challenges with Arabic speakers participating in qualitative research studies. *International Journal of Nursing Studies*, *54*, 150–157. https://doi.org/10.1016/j.ijnurstu.2015.04.010

Alzghool, M. (2009). *Investigating different models for cross-language information retrieval from automatic speech transcripts* [University of Ottawa]. https://ruor.uottawa.ca/handle/10393/29900

Amini, A., Soleimany, A., Schwarting, W., Bhatia, S., & Rus, D. (2019). Uncovering and Mitigating Algorithmic Bias through Learned Latent Structure. *AIES*. http://www.aies-conference.com/wp-content/papers/main/AIES-19_paper_220.pdf

Anderson, M. B. (1999). *Do no harm: How aid can support peace – Or war*. Lynne Rienner Publishers.

Anokwa, Y., Hartung, C., & Brunette, W. (2009). Open source data collection in the developing world. *Computer*, *42*(10), 97–99. https://doi.org/10.1109/MC.2009.328

Ansari, A., & Petras, R. (2018). *Gamayun: The Language Equality Initiative*. Translators Without Borders. https://translatorswithoutborders.org/wp-content/uploads/2018/03/Gamayun-Language-Equality-Initiative-March-2018.pdf

Anthopoulos, L., Reddick, C. G., Giannakidou, I., & Mavridis, N. (2016). Why e-government projects fail? An analysis of the Healthcare.gov website. *Government Information Quarterly*, *33*(1), 161–173. https://doi.org/10.1016/j.giq.2015.07.003

Arksey, H., & O'Malley, L. (2005). Scoping studies: Towards a methodological framework. *International Journal of Social Research Methodology*, *8*(1), 19–32. https://doi.org/10.1080/1364557032000119616

Assal, H., Seng, J., Kurfess, F., Schwarz, E., & Pohl, K. (2011). Semantically-enhanced information extraction. *2011 Aerospace Conference*, 1–14. https://doi.org/10.1109/AERO.2011.5747547

Banatvala, N. (2000). Conflict and health: Public health and humanitarian interventions: developing the evidence base. *BMJ*, *321*(7253), 101–105. https://doi.org/10.1136/bmj.321.7253.101

Bansal, S., Kamper, H., Livescu, K., Lopez, A., & Goldwater, S. (2018). Low-resource speech-to-text translation. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, *2018-Septe*, 1298–1302. https://doi.org/10.21437/Interspeech.2018-1326

Bapna, A., Caswell, I., Kreutzer, J., Firat, O., van Esch, D., Siddhant, A., Niu, M., Baljekar, P., Garcia, X., Macherey, W., Breiner, T., Axelrod, V., Riesa, J., Cao, Y., Chen, M. X., Macherey, K., Krikun, M., Wang, P., Gutkin, A., … Hughes, M. (2022). Building machine

translation systems for the next thousand languages. *ArXiv:2205.03983 [Cs]*. http://arxiv.org/abs/2205.03983

Barman, B., Srivastava, T., Sarma, A., & Nath, C. (2020). Effectiveness of formal training in bioethics of 3 rd semester undergraduate medical students in recognizing bioethical issues and principles in patient care. *Journal of Family Medicine and Primary Care*, *9*(6), 2871. https://doi.org/10.4103/jfmpc.jfmpc_405_20

Barnett, M. N. (2011). *Empire of humanity: A history of humanitarianism*. Cornell University Press.

Barnett, M. N. (2013). Humanitarian Governance. *Annual Review of Political Science*, *16*(1), 379–398. https://doi.org/10.1146/annurev-polisci-012512-083711

Basma, S., Lord, B., Jacks, L. M., Rizk, M., & Scaranelo, A. M. (2011). Error rates in breast imaging reports: Comparison of automatic speech recognition and dictation transcription. *American Journal of Roentgenology*, *197*(4), 923–927. https://doi.org/10.2214/AJR.11.6691

Bass, G. J. (2008). *Freedom's battle: the origins of humanitarian intervention*. Alfred A. Knopf.

Bauer, J.-M., Akakpo, K., Enlund, M., & Passeri, S. (2013). Tracking Vulnerability in Real Time: Mobile Text for Food Security Surveys in Eastern Democratic Republic of Congo. *Africa Policy Journal*, *9*, 36. https://search.proquest.com/docview/1620000489

Beauchamp, T. L., & Childress, J. F. (2019). *Principles of Biomedical Ethics* (7th editio). Oxford University Press.

Belgrave, L. L., & Seide, K. (2018). Grounded Theory Methodology: Principles and Practices. In P. Liamputtong (Ed.), *Handbook of Research Methods in Health Social Sciences* (pp. 1–18). Springer, Singapore. https://doi.org/10.1007/978-981-10-2779-6_84-1

Bempong, N. E., De Castañeda, R. R., Schütte, S., Bolon, I., Keiser, O., Escher, G., & Flahault, A. (2019). Precision Global Health - The case of Ebola: A scoping review. *Journal of Global Health*, *9*(1). https://doi.org/10.7189/jogh.09.010404

Ben-Zeev, D., Schueller, S. M., Begale, M., Duffecy, J., Kane, J. M., & Mohr, D. C. (2015). Strategies for mHealth Research: Lessons from 3 Mobile Intervention Studies. *Administration and Policy in Mental Health and Mental Health Services Research*, *42*(2), 157–167. https://doi.org/10.1007/s10488-014-0556-2

Benelli, P., & Low, T. (2019). Ethical primary research by humanitarian actors. *Forced Migration Review*, *June*, 28–29. https://www.fmreview.org/sites/fmr/files/FMRdownloads/en/ethics/benelli-low.pdf

Betts, A., & Bloom, L. (2014). *Humanitarian Innovation: The State of the Art* (Issue November). UN OCHA. https://www.unocha.org/sites/unocha/files/Humanitarian Innovation The State of the Art_0.pdf

Bharania, R. (2017). *Exploring the Intersection of Technology and Emergencies The GDPR is a unique opportunity to get humanitarian data protection right* . The Digital Responder. https://thedigitalresponder.wordpress.com/2017/12/30/the-gdpr-is-a-unique-opportunity-to-get-humanitarian-data-protection-right/

Biesenthal, C. (2014). Pragmatism. In David Coghlan & Mary Brydon-Miller (Ed.), *The SAGE Encyclopedia of Action Research* (pp. 648–650). SAGE Publications Ltd. https://doi.org/10.4135/9781446294406

Binet, L., & Saulnier, M. (2019). *Médecins Sans Frontières, evolution of an international movement: associative History 1971-2011*. http://associativehistory.msf.org/sites/default/files/MSF Associative history 1971_2011.pdf

Blanchet, K., Ramesh, A., Frison, S., Warren, E., Hossain, M., Smith, J., Knight, A., Post, N., Lewis, C., Woodward, A., Dahab, M., Ruby, A., Sistenich, V., Pantuliano, S., & Roberts, B. (2017). Evidence on public health interventions in humanitarian crises. *The Lancet*, *390*(10109), 2287–2296. https://doi.org/10.1016/S0140-6736(16)30768-1

Bonino, F., Jean, I., & Knox Clarke, P. (2014). *Humanitarian Feedback Mechanisms: Research, Evidence and Guidance*. ALNAP/ODI. https://www.cdacollaborative.org/publication/humanitarian-feedback-mechanisms-research-evidence-and-guidanc/

Bowden, A., & Fox-Rushby, J. A. (2003). A systematic and critical review of the process of translation and adaptation of generic health-related quality of life measures in Africa, Asia, Eastern Europe, the Middle East, South America. *Social Science & Medicine*, *57*, 1289–1306. https://doi.org/10.1016/S0277-9536(02)00503-8

Brenner, N. (2004). *New state spaces : urban governance and the rescaling of statehood*. Oxford University Press.

Broussard, G., Rubenstein, L. S., Robinson, C., Maziak, W., Gilbert, S. Z., & DeCamp, M. (2019). Challenges to ethical obligations and humanitarian principles in conflict settings: a systematic review. *Journal of International Humanitarian Action*, *4*(1). https://doi.org/10.1186/s41018-019-0063-x

Brundage, M., Avin, S., Clark, J., Toner, H., Eckersley, P., Garfinkel, B., Dafoe, A., Scharre, P., Zeitzoff, T., Filar, B., Anderson, H., Roff, H., Allen, G. C., Steinhardt, J., Flynn, C., hÉigeartaigh, S. Ó., Beard, S., Belfield, H., Farquhar, S., … Amodei, D. (2018). *The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation*. http://arxiv.org/abs/1802.07228

Building Markets, & Orange Door Research. (2018). *What is the Point if Nothing Changes: Current Practices and Future Opportunities to Improve Remote Monitoring and Evaluation in Syria*. New York, NY: Building Markets.

Buolamwini, J. (2018). Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. *Proceedings of Machine Learning Research*, *81*, 1–15. http://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf

Burns, R. (2019). New Frontiers of Philanthro-capitalism: Digital Technologies and Humanitarianism. *Antipode*, *51*(4), 1101–1122. https://doi.org/10.1111/anti.12534

Cawthorne, D., & Wynsberghe, A. R. Van. (2019). From HealthDrone to FrugalDrone: Value-Sensitive Design of a Blood Sample Transportation Drone. *International Symposium on Technology and Society, Proceedings*, *2019-Novem*. https://doi.org/10.1109/ISTAS48451.2019.8938019

Chataigner, P. (2017). *Analytical Framework Review report- A framework analysis of analytical frameworks*. OCHA. https://assessments.hpc.tools/km/analytical-framework-review-report

CHS. (2014). *The Core Humanitarian Standard on Quality and Accountability* (First edit). CHS Alliance, Groupe URD and the Sphere Project. http://www.corehumanitarianstandard.org/files/files/Core Humanitarian Standard - English.pdf

Chun Tie, Y., Birks, M., & Francis, K. (2019). Grounded theory research: A design framework for novice researchers. *SAGE Open Medicine*, *7*, 205031211882292. https://doi.org/10.1177/2050312118822927

Cinnamon, J., Jones, S. K., & Adger, W. N. (2016). Evidence and future potential of mobile phone data for disease disaster management. *Geoforum*, *75*, 253–264. https://doi.org/10.1016/j.geoforum.2016.07.019

Clarke, P. K. (2018). *The State of the Humanitarian System 2018*. ALNAP/ODI. https://sohs.alnap.org/help-library/the-state-of-the-humanitarian-system-2018-full-report

Comes, T., Meesters, K., & Torjesen, S. (2019). Making sense of crises: the implications of information asymmetries for resilience and social justice in disaster-ridden communities. *Sustainable and Resilient Infrastructure*, *4*(3), 124–136. https://doi.org/10.1080/23789689.2017.1405653

CommonVoice. (2019). *Languages*. https://voice.mozilla.org/en/languages

Coqui. (2022). *Coqui*. https://coqui.ai/

Corbin, J., & Strauss, A. (2015). *Basics of Qualitative Research* (4th editio). Sage Publications, Inc.

Cornet, V. P., Daley, C., Bolchini, D., Toscos, T., Mirro, M. J., & Holden, R. J. (2019). Patient-centered Design Grounded in User and Clinical Realities: Towards Valid Digital Health. *Proceedings of the International Symposium on Human Factors and Ergonomics in Health Care*, *8*(1), 100–104. https://doi.org/10.1177/2327857919081023

Cornet, V. P., & Holden, R. J. (2018). Systematic review of smartphone-based passive sensing for health and wellbeing. *Journal of Biomedical Informatics*, *77*(October 2017), 120–132. https://doi.org/10.1016/j.jbi.2017.12.008

Cornet, V. P., Toscos, T., Bolchini, D., Ghahari, R. R., Ahmed, R., Daley, C., Mirro, M. J., & Holden, R. J. (2020). Untold stories in user-centered design of mobile health: Practical

challenges and strategies learned from the design and evaluation of an app for older adults with heart failure. *JMIR MHealth and UHealth*, *8*(7), 1–14. https://doi.org/10.2196/17703

Cosgrave, J. (2009). Humanitarian funding and needs assessment. In *Humanitarian Response Index 2008: Donor Accountability in Humanitarian Action* (pp. 75–88). https://doi.org/10.1057/9780230584617

Crawford, K., & Finn, M. (2015). The limits of crisis data: analytical and ethical challenges of using social and mobile data to understand disasters. *GeoJournal*, *80*(4), 491–502. https://doi.org/10.1007/s10708-014-9597-z

Creswell, J. W. (2014). *Research Design* (4th ed.). Sage.

Creswell, J. W. (2018). *Research Design* (5th ed.). Sage.

Cunningham, H., Tablan, V., Roberts, A., & Bontcheva, K. (2013). Getting More Out of Biomedical Documents with GATE's Full Lifecycle Open Source Text Analytics. *PLoS Computational Biology*, *9*(2), e1002854. https://doi.org/10.1371/journal.pcbi.1002854

Currion, P. (2020). The Black Hole of Humanitarian Innovation. *Journal of Humanitarian Affairs*, *1*(3), 42–45. https://doi.org/10.7227/jha.024

Dahdaleh Institute for Global Health Research. (2019, July 15). *Natural Language Processing Project Takes Shape*. https://www.yorku.ca/dighr/natural-language-processing-project-takes-shape/

Darcy, J. (2003). *Measuring humanitarian need: A critical review of needs assessment practice and its influence on resource allocation*. https://www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/4899.pdf

Darcy, J. (2004). *Human Rights and Humanitarian Action: A review of the issues*. https://www.odi.org/publications/1558-human-rights-and-humanitarian-action-review-issues

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly: Management Information Systems*, *13*(3), 319–339. https://doi.org/10.2307/249008

de Waal, A. (1997). *Famine Crimes: Politics & the Disaster Relief Industry in Africa*. Indiana University Press.

DEEP. (2021). *Data Entry and Exploration Platform*.

Dette, R. (2018). Do No Digital Harm: Mitigating Technology Risks in Humanitarian Contexts. In S. Hostettler, S. Najih Besson, & J.-C. Bolay (Eds.), *Technologies for Development* (pp. 13–29). Springer. https://doi.org/10.1007/978-3-319-91068-0_2

Dette, R., Steets, J., & Sagmeister, E. (2016). *Technologies for monitoring in insecure environments*. Global Public Policy Institute. http://www.gppi.net/fileadmin/user_upload/media/pub/2016/SAVE__2016__Toolkit_on_Technologies_for_Monitoring_in_Insecure_Environments.pdf

Deutscher, G. (2010). *Through the Language Glass: Why the World Looks Different in Other Languages*. Henry Hold and Company.

Dew, K. N., Turner, A. M., Choi, Y. K., Bosold, A., & Kirchhoff, K. (2018). Development of machine translation technology for assisting health communication: A systematic review. *Journal of Biomedical Informatics*, *85*(July), 56–67. https://doi.org/10.1016/j.jbi.2018.07.018

Donini, A., & Maxwell, D. (2013). From face-to-face to face-to-screen: Remote management, effectiveness and accountability of humanitarian action in insecure environments. *International Review of the Red Cross*, *95*(890), 383–413. https://doi.org/10.1017/S1816383114000265

Duffield, M. (2008). Global Civil War: The Non-Insured, International Containment and Post-Interventionary Society. *Journal of Refugee Studies*, *21*(2), 145–165. https://doi.org/10.1093/jrs/fem049

Duffield, M. (2016). The resilience of the ruins: towards a critique of digital humanitarianism. *Resilience*, *4*(3), 147–165. https://doi.org/10.1080/21693293.2016.1153772

Earle-Richardson, G., Erlach, E., Walz, V., Baggio, O., Kurnit, M., Camara, C. A., Craig, C., Dios, L. R., Yee, D., Soke, G. N., Voahary, I., & Prue, C. E. (2021). New mixed methods approach for monitoring community perceptions of ebola and response efforts in the democratic Republic of the Congo. *Global Health Science and Practice*, *9*(2), 332–343. https://doi.org/10.9745/GHSP-D-21-00144

Eaton-Lee, J., & Shaughnessy, E. (2021, June 25). *Oxfam's new policy on biometrics explores safe and responsible data practice*. Reliefweb. https://reliefweb.int/report/world/oxfam-s-new-policy-biometrics-explores-safe-and-responsible-data-practice

Eberhard, D. M., Simons, G. F., & Fennig, C. D. (Eds.). (2022). *Ethnologue: Languages of the World* (25th editi). SIL International. https://www.ethnologue.com/a

Eckenwiler, L., Pringle, J., Boulanger, R., & Hunt, M. (2015). Real-time Responsiveness for Ethics Oversight During Disaster Research. *Bioethics*, *29*(9), 653–661. https://doi.org/10.1111/bioe.12193

Elhra, & Humanitarian Health Ethics. (2021). *Ethics for humanitarian innovation: Toolkit*. https://higuide.elrha.org/wp-content/uploads/2021/09/Elrha-Ethics_for_Humanitarian_Innovation-Toolkit-210913.pdf

Elrha. (2022). *Humanitarian Innovation Fund*. https://www.elrha.org/programme/hif/

Emergency Data Science Workshop. (2018). *Natural Language Processing for Humanitarian Survey Work: Challenge Presented by International Rescue Committee*. https://emergencydatascience.org/challenge_nlp_surveys/

Evidence Partners. (2022). *DistillerSR*. https://www.evidencepartners.com

Facebook. (2017, September 14). *A New Center for Crisis Response on Facebook*. Menlo Park: Facebook. https://newsroom.fb.com/news/2017/09/a-new-center-for-crisis-response-on-facebook/

FAO, & WFP. (2022). *Hunger Hotspots: FAO - WFP early warnings October 2022 to January 2023 Outlook*. https://reliefweb.int/report/world/hunger-hotspots-fao-wfp-early-warnings-acute-food-insecurity-october-2022-january-2023-outlook

Figma Inc. (2021). *Figma*. https://www.figma.com/

Finnigan, G., & Farkas, O. (2020). More than Laboratories: Four Decisive Challenges Confronting Humanitarian Innovation. *Journal of Humanitarian Affairs*, *1*(3), 4–13. https://doi.org/10.7227/jha.020

Fiordelli, M., Diviani, N., & Schulz, P. J. (2013). Mapping mhealth research: A decade of evolution. *Journal of Medical Internet Research*, *15*(5). https://doi.org/10.2196/jmir.2430

Floridi, L. (2016). On Human Dignity as a Foundation for the Right to Privacy. In *Philosophy and Technology* (Vol. 29, Issue 4). https://doi.org/10.1007/s13347-016-0220-8

Furlong, P., & Marsh, D. (2010). A Skin not a Sweater: Ontology and Epistemology. In D. Marsh & G. Stoker (Eds.), *Theory and methods in political science*. Palgrave Macmillan.

Furui, S., Nakamura, M., Ichiba, T., & Iwano, K. (n.d.). *Why is the Recognition of Spontaneous Speech so Hard?* Retrieved June 5, 2019, from http://www.furui.cs.titech.ac.jp

Future of Life Institute. (n.d.). *AI Policy Resources*. Retrieved March 5, 2019, from https://futureoflife.org/ai-policy-resources/

Galanis, N. I., Vafiadis, P., Mirzaev, K. G., & Papakostas, G. A. (2021). Machine Learning Meets Natural Language Processing - The Story so Far. *IFIP Advances in Information and Communication Technology*, *627*, 673–686. https://doi.org/10.1007/978-3-030-79150-6_53

Garnerin, M., Rossato, S., & Besacier, L. (2019). Gender representation in French broadcast corpora and its impact on ASR performance. *AI4TV 2019 - Proceedings of the 1st International Workshop on AI for Smart TV Content Production, Access and Delivery, Co-Located with MM 2019*, 3–9. https://doi.org/10.1145/3347449.3357480

Gazi, T. (2020). Data to the rescue: how humanitarian aid NGOs should collect information based on the GDPR. *Journal of International Humanitarian Action*, *5*(1). https://doi.org/10.1186/s41018-020-00078-0

Gerdin, M., Chataigner, P., Tax, L., Kubai, A., & von Schreeb, J. (2014). Does need matter? Needs assessments and decision-making among major humanitarian health agencies. *Disasters*, *38*(3), 451–464. https://doi.org/10.1111/disa.12065

Ghazali, D. A., Guericolas, M., Thys, F., Sarasin, F., Arcos González, P., & Casalino, E. (2018). Climate change impacts on disaster and emergency medicine focusing on mitigation disruptive effects: An international perspective. *International Journal of Environmental Research and Public Health*, *15*(7), 1–13. https://doi.org/10.3390/ijerph15071379

Gibson, D. G., Pereira, A., Farrenkopf, B. A., Labrique, A. B., Pariyo, G. W., & Hyder, A. A. (2017). Mobile Phone Surveys for Collecting Population-Level Estimates in Low- and Middle-Income Countries: A Literature Review. *Journal of Medical Internet Research*, *19*(5), e139. https://doi.org/10.2196/jmir.7428

Global Health Cluster. (2017). *Operational Guidance on Accountability to Affected Populations (AAP)*. https://www.humanitarianresponse.info/system/files/documents/files/cluster_coordination_reference_module_2015_final.pdf

Global Humanitarian Assistance. (2021). *Global Humanitarian Assistance Report 2021*. Development Initiatives. https://devinit.org/documents/1008/Global-Humanitarian-Assistance-Report-2021.pdf

Global Protection Cluster. (2021, August 11). *Protection Analytical Framework*. https://www.globalprotectioncluster.org/2021/08/11/protection-analytical-framework/

Google Cloud. (n.d.). *Language support*. Retrieved April 20, 2019, from https://cloud.google.com/speech-to-text/docs/languages

Gordon, S., & Donini, A. (2015). Romancing principles and human rights: Are humanitarian principles salvageable? *International Review of the Red Cross*, *97*(897–898), 77–109. https://doi.org/10.1017/S1816383115000727

Gould, J. D., & Lewis, C. (1983). Designing for usability-key principles and what designers think. *Conference on Human Factors in Computing Systems - Proceedings*, *28*(3), 50–53. https://doi.org/10.1145/800045.801579

GPPI, & INSPIRE Consortium. (2019). *Methodology to Assess Coordinated Multi-Sector Needs Assessments*. Global Public Policy Institute. https://www.gppi.net/project/assessing-needs-assessments

Gray, L. M., Wong-Wylie, G., Rempel, G. R., & Cook, K. (2020). Expanding qualitative research interviewing strategies: Zoom video communications. *Qualitative Report*, *25*(5), 1292–1301. https://doi.org/10.46743/2160-3715/2020.4212

Greenough, P. G., Chan, J. L., Meier, P., Bateman, L., & Dutta, S. (2009). Applied technologies in humanitarian assistance: Report of the 2009 applied technology working group. *Prehospital and Disaster Medicine*, *24*(SUPPL.2), 2–5. https://doi.org/10.1017/S1049023X00021609

Greenwood, F., Howarth, C., Poole, D. E., Raymond, N. A., & Scarnecchia, D. P. (2017). *The Signal Code: A Human Rights Approach to Information During Crisis*. Cambridge, MA: Harvard Humanitarian Initiative. https://signalcodeorg.files.wordpress.com/2017/01/signalcode_final7.pdf

Groupe URD. (2018). *Compas: Qualité & Redevabilité*. Groupe URD. http://www.urd.org/fr/wp-content/uploads/2018/10/Guide_Compas_FR_V2_mail_091018.pdf

Gu, J., Wang, Y., Chen, Y., Cho, K., & Li, V. O. K. (2018). *Meta-Learning for Low-Resource Neural Machine Translation*. http://arxiv.org/abs/1808.08437

Guba, E., & Lincoln, Y. (2005). Paradigmatic Controversies, Contradictions, and Emerging Confluences. In N. K. Denzin & Y. S. Lincoln (Eds.), *The SAGE Handbook of Qualitative Research* (pp. 191–215). Sage Publications.

Hannun, A., Case, C., Casper, J., Catanzaro, B., Diamos, G., Elsen, E., Prenger, R., Satheesh, S., Sengupta, S., Coates, A., & Ng, A. Y. (2014). *Deep Speech: Scaling up end-to-end speech recognition*. http://arxiv.org/abs/1412.5567

Harvard Humanitarian Initiative. (2011). *Disaster Relief 2.0: The Future of Information Sharing in Humanitarian Emergencies*. https://hhi.harvard.edu/sites/default/files/publications/disaster-relief-2.0.pdf

Haworth, B. T., Bruce, E., Whittaker, J., & Read, R. (2018). The good, the bad, and the uncertain: Contributions of volunteered geographic information to community disaster resilience. *Frontiers in Earth Science*, *6*(November), 1–15. https://doi.org/10.3389/feart.2018.00183

Hayes, P., & Kelly, S. (2018). Distributed morality, privacy, and social media in natural disaster response. *Technology in Society*, *54*, 155. http://ezproxy.library.yorku.ca/login?url=https://search.proquest.com/docview/2113599412?accountid=15182

Hazard, G. C. (1995). Law, Morals, and Ethics. *Southern Illinois University Law Journal*, *19*(3), 447–458. https://heinonline.org/HOL/Page?handle=hein.journals/siulj19&id=477&collection=journals&index=

HHI and ICRC. (2018). *Engaging with people affected by armed conflicts and other situations of violence – Taking stock. Mapping trends. Looking ahead. Recommendations for humanitarian organizations and donors in the digital era.* https://www.icrc.org/en/publication/engaging-people-affected-armed-conflicts-and-other-siutations-violence

Hickman, L. A., Neubert, S., & Reich, K. (Eds.). (2009). *John Dewey between pragmatism and constructivism*. Fordham University Press.

High-Level Expert Group on Artificial Intelligence. (2019a). *A definition of Artificial Intelligence: main capabilities and scientific disciplines*. https://ec.europa.eu/digital-single-market/en/news/definition-artificial-intelligence-main-capabilities-and-scientific-disciplines

High-Level Expert Group on Artificial Intelligence. (2019b). *Ethics guidelines for trustworthy AI*. https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=58477

Hilhorst, D., & Schmiemann, N. (2002). Development in Practice Humanitarian principles and organisational culture: Everyday practice in Médecins Sans Frontières-Holland. *Development in Practice*, *12*(3–4), 490–500. https://doi.org/10.1080/0961450220149834

Holden, R. J., & Boustani, M. A. (2021). Design and Usability Methods. In R. J. Holden & R. S. Valdez (Eds.), *The Patient Factor* (1st ed.). CRC Press.

Holden, R. J., Boustani, M. A., & Azar, J. (2021). Agile Innovation to transform healthcare: Innovating in complex adaptive systems is an everyday process, not a light bulb event. *BMJ Innovations*, *7*(2), 499–505. https://doi.org/10.1136/bmjinnov-2020-000574

Humanitarian Accountability Partnership. (2010). *The 2010 HAP Standard in Accountability and Quality Management* (Second edition). HAP International. https://reliefweb.int/sites/reliefweb.int/files/resources/260222E062A32EF085257829005BD6FE-Full_report.pdf

Humanitarian Grand Challenges. (2022). *Creating Hope in Conflict: A Humanitarian Grand Challenge*. https://humanitariangrandchallenge.org/

Humanitarian Outcomes/Aid Worker Security Database. (2022). *Aid Worker Security Report: Figures at a glance*. Humanitarian Outcomes.

https://www.humanitarianoutcomes.org/sites/default/files/publications/awsd_figures_2022.pdf

Humanitarian Outcomes. (2021). *Aid Worker Security Database*. https://aidworkersecurity.org/incidents/report

Humanitarian Programme Cycle. (2021). *Humanitarian Response Plan Guidance*. OCHA.

Hunt, M., Pringle, J., Christen, M., Eckenwiler, L., Schwartz, L., & Davé, A. (2016). Ethics of emergent information and communication technology applications in humanitarian medical assistance. *International Health*, *8*(4), 239–245. https://doi.org/10.1093/inthealth/ihw028

Hunt, M. R. (2008). Ethics beyond borders: How health professionals experience ethics in humanitarian assistance and development work. *Developing World Bioethics*, *8*(2), 59–69. https://doi.org/10.1111/j.1471-8847.2006.00153.x

IASC. (1999). *Protection of Internally Displaced Persons*. https://interagencystandingcommittee.org/system/files/legacy_files/protection_of_internally_displaced_persons_inter_agency_standing_committee_policy_paper_0.pdf

IASC. (2015). Multi-Sector Initial Rapid Assessment Guidance. In *Geneva: Inter-Agency Standing Committee*. https://interagencystandingcommittee.org/system/files/mira_2015_final.pdf

IASC. (2019a). *Ensuring Data and Analysis is Useful and Usable for Response - Tools*. Inter-Agency Standing Committee. https://interagencystandingcommittee.org/system/files/ensuring_data_and_analysis_is_useful_and_usable_for_response_-_tools.pdf

IASC. (2019b). *Workstream 5: Tools and guidance for advancing coordinated needs assessment and analysis through the Grand Bargain*. Inter-Agency Standing Committee.

https://interagencystandingcommittee.org/improve-joint-and-impartial-needs-assessments/workstream-5-tools-and-guidance-advancing-coordinated

IASC. (2021). *Operational Guidance: Data Responsibility in Humanitarian Action*. Inter-Agency Standing Committee. https://interagencystandingcommittee.org/operational-response/iasc-operational-guidance-data-responsibility-humanitarian-action

ICRC. (2015). *Ethical Principles of Health Care in Times of Armed Conflict and Other Emergencies*. International Committee of the Red Cross. https://www.icrc.org/en/document/common-ethical-principles-health-care-conflict-and-other-emergencies

ICRC. (2016). *Protracted conflict and humanitarian action | ICRC*. International Committee of the Red Cross. https://www.icrc.org/en/document/protracted-conflict-and-humanitarian-action

IDMC. (2021). *2021 Global Report on Internal Displacement*. Internal Displacement Monitoring Centre. https://www.internal-displacement.org/global-report/grid2021/

IDMC. (2022). *2022 Global Report on Internal Displacement*. Internal Displacement Monitoring Centre. https://www.internal-displacement.org/sites/default/files/publications/documents/IDMC_GRID_2022_LR.pdf

IFRC. (1994). *The Code of Conduct for the International Red Cross and Red Crescent Movement and Non Governmental Organisations (NGOs) in Disaster Relief*. http://www.ifrc.org/Global/Publications/disasters/code-of-conduct/code-english.pdf

IFRC. (2006). *Statutes of the International Red Cross and Red Crescent Movement*. 25th International Conference of the Red Cross. https://www.ifrc.org/PageFiles/88219/statutes-en-a5.pdf

International Committee of the Red Cross. (2017). *Handbook on Data Protection in Humanitarian Action* (C. Kuner & M. Marelli (Eds.); 2017th ed.). International Committee of the Red Cross. https://shop.icrc.org/e-books/handbook-on-data-protection-in-humanitarian-action.html

International Committee of the Red Cross. (2019). *Accountability to Affected People Institutional Framework*. https://www.icrc.org/en/publication/accountability-affected-people-institutional-framework

International Committee of the Red Cross. (2020). *Handbook on Data Protection in Humanitarian Action (2nd Edition)*. https://www.icrc.org/en/data-protection-humanitarian-action-handbook

International Committee of the Red Cross and Privacy International. (2018). *The Humanitarian Metadata Problem: "Doing No Harm" in the Digital Era*. https://privacyinternational.org/sites/default/files/2018-12/The Humanitarian Metadata Problem - Doing No Harm in the Digital Era.pdf

International Office for Migration. (2021, November 23). *Displacement Tracking Matrix*. https://dtm.iom.int/

International Organization for Standardization. (2007). *Codes for the representation of names of languages — Part 3: Alpha-3 code for comprehensive coverage of languages*. ISO 639-3:2007.

International Organization for Standardization. (2018, March). *Ergonomics of human-system interaction — Part 11: Usability: Definitions and concepts*. ISO 9241-11:2018(En). https://kebs.isolutions.iso.org/obp/ui#iso:std:iso:9241:-11:ed-2:v1:en

International Rescue Committee. (2018). *Obtaining Meaningful Informed Consent*. https://www.principlesinpractice.info/help-library/irc-research-toolkit-obtaining-meaningful-informed-consent

IOM. (2021). *Counter-trafficking in Emergencies: Information Management Guide*. International Officer for Migration. https://publications.iom.int/books/countertrafficking-emergencies-information-management-guide

IPCC. (2022). Summary for Policymakers. In H. O. Pörtner, D. C. Roberts, M. Tignor, E. S. Poloczanska, K. Mintenbeck, A. Alegría, M. Craig, S. Langsdorf, S. Löschke, V. Möller, A. Okem, & B. Rama (Eds.), *Climate Change 2022: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (p. In Press). Cambridge University Press. https://www.ipcc.ch/report/ar6/wg2/downloads/report/IPCC_AR6_WGII_SummaryForPolicymakers.pdf

ITU. (2018). *Measuring the Information Society Report 2018*. https://www.itu.int/en/ITU-D/Statistics/Documents/publications/misr2018/MISR-2018-Vol-1-E.pdf

Jackson, S., & Aakhus, M. (2014). Becoming More Reflective about the Role of Design in Communication. *Journal of Applied Communication Research*, *42*(2), 125–134. https://doi.org/10.1080/00909882.2014.882009

Jacobsen, K. L. (2015a). Experimentation in humanitarian locations: UNHCR and biometric registration of Afghan refugees. *Security Dialogue*, *46*(2), 144–164. https://doi.org/10.1177/0967010614552545

Jacobsen, K. L. (2015b). *The politics of humanitarian technology : good intentions, unintended consequences and insecurity*. Routledge.

Jacobsen, K. L. (2017). On Humanitarian Refugee Biometrics and New Forms of Intervention. *Journal of Intervention and Statebuilding*, *11*(4), 529–551. https://doi.org/10.1080/17502977.2017.1347856

Klein, G., Kim, Y., Deng, Y., Senellart, J., & Rush, A. M. (2017). *OpenNMT: Open-Source Toolkit for Neural Machine Translation*. http://arxiv.org/abs/1701.02810

Kobo. (2022). *About the Kobo Organization*. https://www.kobotoolbox.org/about-us/the-organization/

Kobo. (2022). *KoBoToolbox* (2.021.45). https://www.kobotoolbox.org

KoboToolbox. (2022a). *kobotoolbox/collect*. https://github.com/kobotoolbox/collect

KoboToolbox. (2022b). *kobotoolbox/kpi*. https://github.com/kobotoolbox/kpi/

Koehn, P., & Knowles, R. (2017). Six Challenges for Neural Machine Translation. *Proceedings Ofthe First Workshop on Neural Machine Translation*, 28–39. https://doi.org/10.18653/v1/w17-3204

Koenecke, A., Nam, A., Lake, E., Nudell, J., Quartey, M., Mengesha, Z., Toups, C., Rickford, J. R., Jurafsky, D., & Goel, S. (2020). Racial disparities in automated speech recognition. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(14), 7684–7689. https://doi.org/10.1073/pnas.1915768117

Krebs, H. B. (2014). *Responsibility, legitimacy, morality: Chinese humanitarianism in historical perspective*. ODI/HPG. https://cdn.odi.org/media/documents/9139.pdf

Kreutzer, T. (2009). *Internet and Online Media Usage on Mobile Phones among Low-Income Urban Youth in Cape Town*. Presented at International Communication Association Conference. http://tinokreutzer.org/mobile/InternetOnlineMediaUsage(ICA).pdf

Kreutzer, T. (2019). *Create automatic background recording of interview in Collect*. https://forum.getodk.org/t/create-automatic-background-recording-of-interview-in-collect/18601/

Kreutzer, T. (2021a). *Ethical Implications Related to Processing of Personal Data in Humanitarian Crises: A Scoping Review*.

Kreutzer, T. (2021b). *Overcoming Blind Spots: Constraints and Solutions Related to Qualitative Interview Methods in Humanitarian Needs Assessments*.

Kreutzer, T., Vinck, P., Pham, P., An, A., Appel, L., Deluca, E., Tang, G., Alzghool, M., Hachhethu, K., Morris, B., Crowley, J., & Orbinski, J. (2020). Improving Humanitarian Needs Assessments through Natural Language Processing. *IBM Journal of Research and Development*, *64*(1). https://doi.org/10.1147/JRD.2019.2947014

Krishnaraj, G., Hunt, M., Schwartz, L., O'Mathúna, D., Pringle, J., Okhowat, A., Kapiriri, L., Brahimi, L., & Yantzi, R. (2021). *Ethics for Humanitarian Innovation: Background Paper*. Elrha. https://higuide.elrha.org/wp-content/uploads/2021/09/Elrha-Ethics_for_Humanitarian_Innovation-Background_Paper-210909.pdf

Kushniruk, A., & Nøhr, C. (2016). Participatory design, user involvement and health IT evaluation. *Evidence-Based Health Informatics: Promoting Safety and Efficiency through Scientific Methods and Ethical Policy*, *222*, 139–151. https://doi.org/10.3233/978-1-61499-635-4-139

Lamanna, C., Hachhethu, K., Chesterman, S., Singhal, G., Mwongela, B., Ng'endo, M., Passeri, S., Farhikhtah, A., Kadiyala, S., Bauer, J.-M., & Rosenstock, T. S. (2019). Strengths and limitations of computer assisted telephone interviews (CATI) for nutrition data collection in rural Kenya. *PLoS ONE*, *14*(1), e0210050. https://doi.org/10.1371/journal.pone.0210050

Langhaug, L. F., Cheung, Y. B., Pascoe, S. J. S., Chirawu, P., Woelk, G., Hayes, R. J., & Cowan, F. M. (2011). How you ask really matters: Randomised comparison of four sexual behaviour questionnaire delivery modes in Zimbabwean youth. *Sexually Transmitted Infections*, *87*(2), 165–173. https://doi.org/10.1136/sti.2009.037374

Latif, S., Qayyum, A., Usama, M., Qadir, J., Zwitter, A., & Shahzad, M. (2019). Caveat Emptor: The Risks of Using Big Data for Human Development. *IEEE Technology & Society Magazine*, *38*(3), 82–90. http://10.0.4.85/MTS.2019.2930273

Leaning, J. (1999). Medicine and international humanitarian law. Law provides norms that must guide doctors in war and peace. *BMJ (Clinical Research Ed.)*, *319*(7207), 393–394. http://www.ncbi.nlm.nih.gov/pubmed/10445905

Leo, B., Morello, R., Mellon, J., Peixoto, T., & Davenport, S. (2015). Do Mobile Phone Surveys Work in Poor Countries? *CGD Working Paper*, *398*. https://doi.org/10.2139/ssrn.2597885

Lev Aretz, Y. (2019). Data Philanthropy. *SSRN Electronic Journal*, *70*(6). https://doi.org/10.2139/ssrn.3320798

Levac, D., Colquhoun, H., & O'Brien, K. K. (2010). Scoping studies: Advancing the methodology. *Implementation Science*, *5*(1). https://doi.org/10.1186/1748-5908-5-69

Lischer, S. K. (2003). Collateral Damage: Humanitarian Assistance as a Cause of Conflict. *International Security*, *28*(1), 79–109. https://doi.org/10.1162/016228803322427983

Loper, E., & Bird, S. (2002, May 17). NLTK: The Natural Language Toolkit. *Proceedings of the ACL Workshop on Effective Tools and Methodologies for Teaching Natural Language Processing and Computational Linguistics*. http://arxiv.org/abs/cs/0205028

Madon, S., & Schoemaker, E. (2019). Reimagining Refugee Identity Systems: A Sociological Approach. *Information and Communication Technologies for Development: Strengthening Southern-Driven Cooperation as a Catalyst for Ict4d, Pt I*.

Makhoul, J., Chehab, R. F., Shaito, Z., & Sibai, A. M. (2018). A scoping review of reporting "Ethical Research Practices" in research conducted among refugees and war-affected populations in the Arab world. *BMC Medical Ethics*, *19*(1), 1–9. https://doi.org/10.1186/s12910-018-0277-2

Marco, C. A., Lu, D. W., Stettner, E., Sokolove, P. E., Ufberg, J. W., & Noeller, T. P. (2011). Ethics curriculum for emergency medicine graduate medical education. *Journal of Emergency Medicine*, *40*(5), 550–556. https://doi.org/10.1016/j.jemermed.2010.05.076

Martin-Shields, C. (2013). The Technologists Dilemma: Ethical Challenges of Using Crowdsourcing Technology in Conflict and Disaster-Affected Regions. *Georgetown Journal of International Affairs*, *14*(2), 157–163. http://www.jstor.com/stable/43134422

Martínez, J. C., & Eng, B. (2016). The unintended consequences of emergency food aid: Neutrality, sovereignty and politics in the Syrian civil war, 2012-15. *International Affairs*, *92*(1), 153–173. https://doi.org/10.1111/1468-2346.12502

Martínez, L. L. (2016). Pueblos indígenas y acción humanitaria en América Latina y el Caribe. In I. Mosel, C. Bennett, & H. Krebs (Eds.), *Aproximaciones a la historia del humanitarismo en América Latina y el Caribe*. ODI/HPG. https://cdn.odi.org/media/documents/Aproximaciones_a_la_historia_del_humanitarismo_en_América_Latina_y_el_Caribe.pdf

MAXQDA 2020. (2019). VERBI Software. https://www.maxqda.com

Maya Hasan, M. (2018). *The language lesson: What we've learned about communicating with Rohingya refugees*. https://translatorswithoutborders.org/wp-content/uploads/2018/12/TWB_Bangladesh_Comprehension_Study_Nov2018.pdf

McCallum, A. K. (2002). *MALLET: A Machine Learning for Language Toolkit*. http://mallet.cs.umass.edu/

McCaslin, M. L. (2012). Pragmatism. In L. M. Given (Ed.), *The SAGE Encyclopedia of Qualitative Research Methods* (pp. 672–675). SAGE Publications, Inc. https://doi.org/10.4135/9781412963909

McNamara, A., Smith, J., & Murphy-Hill, E. (2018). *Does ACM's code of ethics change ethical decision making in software development?* 729–733. https://doi.org/10.1145/3236024.3264833

Meier, P. (2011). New information technologies and their impact on the humanitarian sector. *International Review of the Red Cross*, *93*(884). https://doi.org/10.1017/S1816383112000318

Meier, P. (2014). *Digital humanitarians : how big data is changing the face of humanitarian response*. CRC Press.

Mesmar, S., Talhouk, R., Akik, C., Olivier, P., Elhajj, I. H., Elbassuoni, S., Armoush, S., Kalot, J., Balaam, M., Germani, A., & Ghattas, H. (2016). The impact of digital technology on health of populations affected by humanitarian crises: Recent innovations and current gaps. *Journal of Public Health Policy*, *37*(2), S167–S200. https://doi.org/10.1057/s41271-016-0040-1

Microsoft. (2018). *Microsoft reaches a historic milestone, using AI to match human performance in translating news from Chinese to English - Microsoft News Center Hong Kong*.

https://news.microsoft.com/en-hk/2018/03/15/microsoft-reaches-a-historic-milestone-using-ai-to-match-human-performance-in-translating-news-from-chinese-to-english/

Mihalcea, R. (2004). Graph-based ranking algorithms for sentence extraction, applied to text summarization. *Proceedings of the ACL 2004 on Interactive Poster and Demonstration Sessions*. https://doi.org/10.3115/1219044.1219064

Mock, N., Morrow, N., & Papendieck, A. (2013). From complexity to food security decision-support: Novel methods of assessment and their role in enhancing the timeliness and relevance of food and nutrition security information. *Global Food Security*, *2*(1), 41–49. https://doi.org/10.1016/j.gfs.2012.11.007

Molnar, P., & Gill, L. (2018). *Bots at the Gate: A human rights analysis of automated decision-making in Canada's immigration and refugee system*. University of Toronto. https://ihrp.law.utoronto.ca/

Montreal declaration. (2018). *Montreal declaration for a responsible development of artificial intelligence 2018*. https://www.montrealdeclaration-responsibleai.com/the-declaration

Morgan, D. L. (2014). *Integrating Qualitative and Quantitative Methods: A Pragmatic Approach*. SAGE. https://doi.org/10.4135/9781544304533

Morrow, N., Mock, N., Bauer, J. M., & Browning, J. (2016). Knowing just in time: Use cases for mobile surveys in the humanitarian world. *Procedia Engineering*, *159*(June), 210–216. https://doi.org/10.1016/j.proeng.2016.08.163

Moussa, J. (2014). *Ancient origins, modern actors: Defining Arabic meanings of humanitarianism*. ODI/HPG. https://cdn.odi.org/media/documents/9290.pdf

Mozilla. (2022a). *Common Voice*. https://commonvoice.mozilla.org/en

Mozilla. (2022b). *mozilla/DeepSpeech*. https://github.com/mozilla/DeepSpeech

MSF. (n.d.). *Independent, impartial, neutral*. Retrieved January 4, 2021, from https://www.doctorswithoutborders.org/who-we-are/principles/independence

MSF. (2019). *The Data Protection Course (OCP)*. MSF Tembo. https://tembo.msf.org/course/info.php?id=265

MSF. (2022). *Financial Report 2021*. https://www.msf.org/sites/default/files/2022-05/MSF_Financial_Report_2021_FINAL provisional.pdf

MSF Ethics Review Board. (2013). *MSF Research Ethics Framework - Guidance Document*. http://hdl.handle.net/10144/305288

Narang, N. (2015). Assisting Uncertainty: How Humanitarian Aid can Inadvertently Prolong Civil War. *International Studies Quarterly*, *59*(1), 184–195. https://doi.org/10.1111/isqu.12151

National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. (1978). *Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research*. Department of Health, Education, and Welfare. http://videocast.nih.gov/pdf/ohrp_belmont_report.pdf

Neubert, S. (2001). *Pragmatism and Constructivism in Contemporary Philosophical Discourse*. http://www.hwf.uni-koeln.de/data/dewey/File/Neubert_Pragmatism_Constructivism.pdf

O'Brien, B. C., Harris, I. B., Beckman, T. J., Reed, D. A., & Cook, D. A. (2014). Standards for reporting qualitative research: A synthesis of recommendations. *Academic Medicine*, *89*(9), 1245–1251. https://doi.org/10.1097/ACM.0000000000000388

Obrecht, A., & Warner, A. T. (20116). *More than just luck : Innovation in humanitarian action*. Humanitarian Innovation Fund and ALNAP.

OCHA. (n.d.). *Financial Tracking Service*. Retrieved February 1, 2020, from https://fts.unocha.org/

OCHA. (2012). *Humanitarian Principles*. UN Office for the Coordination of Humanitarian Affairs. https://www.unocha.org/sites/dms/Documents/OOM-humanitarianprinciples_eng_June12.pdf

OCHA. (2015). World Humanitarian Data and Trends 2015. In *United Nations Office for the Coordination of Humanitarian Affairs*. https://www.unocha.org/sites/dms/Documents/WHDT2015_2Dec.pdf

OCHA. (2018). *Global Humanitarian Overview 2019*. Office for the Coordination of Humanitarian Affairs. https://www.unocha.org/sites/unocha/files/GHO2019.pdf

OCHA. (2019). *Data responsibility guidelines*. UN Office for the Coordination of Humanitarian Affairs. https://centre.humdata.org/wp-content/uploads/2019/03/OCHA-DR-Guidelines-working-draft-032019.pdf

OCHA. (2021a). *Global Humanitarian Overview 2021*. Office for the Coordination of Humanitarian Affairs. https://doi.org/10.18356/9789214030751

OCHA. (2021b). *Global Humanitarian Overview 2022*. Office for the Coordination of Humanitarian Affairs. https://doi.org/10.18356/9789210012423

OCHA. (2021c). *Joint intersectoral analysis framework 1.1*. UNOCHA. https://doi.org/10.18356/9789214030751c020

OCHA. (2022). *Global Humanitarian Overview 2023*. Office for the Coordination of Humanitarian Affairs. https://reliefweb.int/report/world/global-humanitarian-overview-2023-enaresfr

OCHA. (2023). *Financial Tracking Service*. https://fts.unocha.org

Recommendation of the Council on Artificial Intelligence, Pub. L. No. OECD/LEGAL/0449 (2019). https://doi.org/10.1787/eedfee77-en

Okular Analytics. (2021). *Technical Review of 2020 MSNAs in Protracted Crises*. IMPACT Initiatives. https://www.reachresourcecentre.info/wp-content/uploads/2021/07/Okular_Report_MSNAs2020ExternalReview_July2021.pdf

Open Source Initiative. (2007, March 22). *The Open Source Definition*.

Opensource.com. (n.d.). *What is open source?* Retrieved May 1, 2022, from https://opensource.com/resources/what-open-source

Orbinski, J. (1994). Health Impact of Rights Violations in Conf1ict Zones and the Humanitarian Imperative: Rwanda-A Case Study. *Refuge: Canada's Journal on Refugees*, *14*(5), 5–8. https://doi.org/10.25071/1920-7336.21827

Orbinski, J. (2009). *An Imperfect Offering: Humanitarian Action for the Twenty-First Century*. Walker & Co.

Otter, D. W., Medina, J. R., & Kalita, J. K. (2021). A Survey of the Usages of Deep Learning for Natural Language Processing. *IEEE Transactions on Neural Networks and Learning Systems*, *32*(2), 604–624. https://doi.org/10.1109/TNNLS.2020.2979670

Ozuru, Y., Briner, S., Kurby, C. A., & McNamara, D. S. (2013). Comparing comprehension measured by multiple-choice and open-ended questions. *Canadian Journal of Experimental Psychology*, *67*(3), 215–227. https://doi.org/10.1037/a0032918

Pal, N. E., Eckenwiler, L., Hyppolite, S.-R., Pringle, J., Chung, R., & Hunt, M. (2019). Ethical considerations for closing humanitarian projects: a scoping review. *Journal of International Humanitarian Action*, *4*(1). https://doi.org/10.1186/s41018-019-0064-9

Parker, B. (2019, February 5). *New UN deal with data mining firm Palantir raises protection concerns*. The New Humanitarian. https://www.irinnews.org/news/2019/02/05/un-palantir-deal-data-mining-protection-concerns-wfp

Parsa-Parsi, R. W. (2022). The International Code of Medical Ethics of the World Medical Association. In *JAMA* (Vol. 328, Issue 20, pp. 2018–2021). American Medical Association. https://doi.org/10.1001/jama.2022.19697

Paul, A., Jolley, C., & Anthony, A. (2018). *Reflecting the Past, Shaping the Future: Making AI Work for International Development*. USAID. https://www.usaid.gov/digital-development/machine-learning/AI-ML-in-development

Paul, S., & Sosale, S. (2020). Witnessing a Disaster: Public Use of Digital Technologies in the 2015 South Indian Floods. *Digital Journalism*, *8*(1), 15–31. https://doi.org/10.1080/21670811.2019.1636693

People in Aid. (2003). *Code of Good Practice in the management and support of aid personnel* (Second edi). https://reliefweb.int/sites/reliefweb.int/files/resources/DA4AA643ACE54425C1256C6B003B903A-pia-code.pdf

Perakslis, E. D. (2018). Using digital health to enable ethical health research in conflict and other humanitarian settings. *Conflict and Health*, *12*(1), 23. https://doi.org/10.1186/s13031-018-0163-z

Peters, M. D. J., Godfrey, C. M., Khalil, H., McInerney, P., Parker, D., & Soares, C. B. (2015). Guidance for conducting systematic scoping reviews. *International Journal of Evidence-Based Healthcare*, *13*(3), 141–146. https://doi.org/10.1097/XEB.0000000000000050

Peters, M. D. J., Godfrey, C., Mcinerney, P., Baldini Soares, C., Khalil, H., & Parker, D. (2017). Chapter 11: Scoping reviews. In E. Aromataris & Z. Munn (Eds.), *Joana Briggs Institute Reviewer's Manual* (Issue September, pp. 6–24). https://reviewersmanual.joannabriggs.org/

Pham, K. T., Sattigeri, P., Dhurandhar, A., Jacob, A. C., Vukovic, M., Chataigner, P., Freire, J., Mojsilovic, A., & Varshney, K. R. (2017). Real-time understanding of humanitarian crises via targeted information retrieval. *IBM Journal of Research and Development*, *61*(6), 7:1-7:12. https://doi.org/10.1147/JRD.2017.2722799

Pham, P. N., & Vinck, P. (2012). Technology fusion and their implications for conflict early warning systems, public health, and human rights. *Health and Human Rights*, *14*(2), 106–117. http://www.jstor.org/stable/healhumarigh.14.2.106

Phillips, J. (2018). Risk in a digital age: understanding risk in virtual networks through digital response networks (DRNs). *International Development Planning Review*, *40*(3), 239–272. https://doi.org/10.3828/idpr.2018.18

Pictet, J. (1979). *The fundamental principles of the Red Cross*. http://www.ifrc.org/PageFiles/40669/Pictet Commentary.pdf

Polman, L. (2010). *The Crisis Caravan: What's Wrong with Humanitarian Aid*. Metropolitan Books.

Poole, D., Latonero, M., & Berens, J. (2017). *Refugee Connectivity: A Survey of Mobile Phones, Mental Health, and Privacy at a Syrian Refugee Camp in Greece*. Cambridge, MA: Harvard Humanitarian Initiative and New York, NY: Data & Society Research Institute.

Popping, R. (2015). Analyzing Open-ended Questions by Means of Text Analysis Procedures. *BMS Bulletin of Sociological Methodology/ Bulletin de Methodologie Sociologique*, *128*(1), 23–39. https://doi.org/10.1177/0759106315597389

Ramalingam, B., Scriven, K., & Foley, C. (2009). *Innovations in international humanitarian action: ALNAP's 8th review of humanitarian action*. ALNAP. https://www.alnap.org/help-library/innovations-in-international-humanitarian-action-alnaps-8th-review-of-humanitarian

Raymond, N. A. (2017). Beyond "Do No Harm" and Individual Consent: Reckoning with the Emerging Ethical Challenges of Civil Society's Use of Data. In B. Taylor, L and Floridi, L and VanDerSloot (Ed.), *Group Privacy: New Challenges of Data Technologies* (Vol. 126, pp. 67–82). Springer. https://doi.org/10.1007/978-3-319-46608-8_4

Raymond, N. A., Al Achkar, Z., Verhulst, S., & Berens, J. (2016). Building data responsibility into humanitarian action. *OCHA Policy and Studies Series*, *018*, 18. https://docs.unocha.org/sites/dms/Documents/TB18_Data Responsibility_Online.pdf

Raymond, N. A., & Card, B. (2015). *Applying Humanitarian Principles to Current Uses of Information Communication Technologies: Gaps in Doctrine and Challenges to Practice*. Harvard Humanitarian Initiative. https://hhi.harvard.edu/sites/default/files/publications/signal_program_humanitarian_principles_white_paper.pdf

Redmond, A. D. (2005). Needs assessment of humanitarian crises. *Bmj*, *330*(7503), 1320–1322. https://doi.org/10.1136/bmj.330.7503.1320

Roberts, C. h., Brindle, H., Rogers, N. T., Eggo, R. M., Enria, L., & Lees, S. (2021). Vaccine Confidence and Hesitancy at the Start of COVID-19 Vaccine Deployment in the UK: An Embedded Mixed-Methods Study. *Frontiers in Public Health*, *9*(November). https://doi.org/10.3389/fpubh.2021.745630

Roberts, T., & Faith, B. (2021). *Digital Aid: understanding the digital challenges facing humanitarian assistance*. *March*. https://doi.org/10.19088/IDS.2021.030

Robinson, A., & Obrecht, A. (2016). *Using mobile voice technology to improve the collection of food security data: WFP's mobile Vulnerability Analysis and Mapping*. ODI/ALNAP. https://www.alnap.org/system/files/content/resource/files/main/alnap-wfp-mvam-case-study-2016.pdf

Russell, S. J., & Norvig, P. (2009). *Artificial Intelligence: A Modern Approach* (3rd editio). Prentice Hall.

Sajad Mousavi, S., Schukat, M., & Howley, E. (2018). Deep Reinforcement Learning: An Overview. *SAI Intelligent Systems Conference (IntelliSys)*. https://arxiv.org/pdf/1806.08894.pdf

Sandvik, K. B., Gabrielsen Jumbert, M., Karlsrud, J., & Kaufmann, M. (2014). Humanitarian technology: a critical research agenda. *International Review of the Red Cross*, *96*(893), 219–242. https://doi.org/10.1017/s1816383114000344

Sandvik, K. B., Jacobsen, K. L., & McDonald, S. M. (2017). Do no harm: A taxonomy of the challenges of humanitarian experimentation. *International Review of the Red Cross*, *99*(904), 319–344. https://doi.org/10.1017/S181638311700042X

Sandvik, K. B., Jumbert, M. G., Karlsrud, J., & Kaufmann, M. (2014). Humanitarian technology: a critical research agenda. *International Review of the Red Cross*, *96*(893), 219. https://doi.org/http://dx.doi.org/10.1017/S1816383114000344

Sandvik, K. B., & Raymond, N. A. (2017). Beyond the Protective Effect: Towards a Theory of Harm for Information Communication Technologies in Mass Atrocity Response. *Genocide Studies and Prevention: An International Journal*, *11*(1), 9–24. https://doi.org/10.5038/1911-9933.11.1.1454

Sapkota, P. P., & Siddiqi, K. (2019). Is Ubiquitous Technology for Needs Data Management a Game Changer in Humanitarian Arena? *International Journal of Information Systems for Crisis Response and Management*, *11*(1), 83–97. https://doi.org/10.4018/IJISCRAM.2019010105

Sarı, L., Hasegawa-Johnson, M., & Yoo, C. D. (2021). Counterfactually Fair Automatic Speech Recognition. *IEEE/ACM Transactions on Audio Speech and Language Processing*, *29*, 3515–3525. https://doi.org/10.1109/TASLP.2021.3126949

Savoldelli, A., Codagnone, C., & Misuraca, G. (2014). Understanding the e-government paradox: Learning from literature and practice on barriers to adoption. *Government Information Quarterly*, *31*(SUPPL.1), S63–S71. https://doi.org/10.1016/j.giq.2014.01.008

Schofield, G., Dittborn, M., Selman, L. E., & Huxtable, R. (2021). Defining ethical challenge(s) in healthcare research: a rapid review. *BMC Medical Ethics*, *22*(1), 1–17. https://doi.org/10.1186/s12910-021-00700-9

Scholte, J. A. (2014). Reinventing global democracy. *European Journal of International Relations*, *20*(1), 3–28. https://doi.org/10.1177/1354066111436237

Schopper, D., Dawson, A., Upshur, R., Ahmad, A., Jesani, A., Ravinetto, R., Segelid, M. J., Sheel, S., & Singh, J. (2015). Innovations in research ethics governance in humanitarian settings. *BMC Medical Ethics*, *16*(10), 1–12. https://doi.org/10.1186/s12910-015-0002-3

Schwartz, L., Hunt, M., Sinding, C., Elit, L., Redwood-Campbell, L., Adelson, N., & De Laat, S. (2012). Models for humanitarian health care ethics. *Public Health Ethics*, *5*(1), 81–90. https://doi.org/10.1093/phe/phs005

Scott-Smith, T. (2016). Humanitarian neophilia: the ‘innovation turn’ and its implications. *Third World Quarterly*, *37*(12), 2229–2251. https://doi.org/10.1080/01436597.2016.1176856

Sean McDonald. (2016). *Ebola: A Big Data Disaster* (2016.01; CIS Papers). https://cis-india.org/papers/ebola-a-big-data-disaster

Sheather, J., Jobanputra, K., Schopper, D., Pringle, J., Venis, S., Wong, S., & Vincent-Smith, R. (2016). A Médecins Sans Frontières Ethics Framework for Humanitarian Innovation. *PLoS Medicine*, *13*(9), 1–10. https://doi.org/10.1371/journal.pmed.1002111

Shoemaker, E., Kristinsdottir, G. S., Ahuja, T., Baslan, D., Pon, B., Currion, P., Gumisizira, P., & Dell, N. (2019). Identity at the margins: Examining refugee experiences with digital identity systems in Lebanon, Jordan, and Uganda. *COMPASS 2019 - Proceedings of the 2019 Conference on Computing and Sustainable Societies*, 206–217. https://doi.org/10.1145/3314344.3332486

Simms, B., & Trim, D. J. B. (2011). *Humanitarian Intervention: A History*. Cambridge University Press.

Simon, H. A. (1996). *The sciences of the artificial* (3rd ed.). MIT Press.

Singer, P. (2021). Ethics. In *Encyclopedia Britannica*.

Singh, S. (2018). *Natural Language Processing for Information Extraction*. http://arxiv.org/abs/1807.02383

Slim, H. (2015). *Humanitarian Ethics: A Guide to the Morality of Aid in War and Disaster*. Oxford University Press.

Smith, A., Pringle, J., & Hunt, M. (2020). Value-Sensitive Design for Humanitarian Action: Integrating Ethical Analysis for Information and Communication Technology Innovations. In D. Messelken & D. Winkler (Eds.), *Ethics of Medical Innovation, Experimentation, and Enhancement in Military and Humanitarian Contexts* (pp. 105–124). Springer. https://doi.org/10.1007/978-3-030-36319-2_7

Sphere Association. (2018). *The Sphere Handbook: Humanitarian Charter and Minimum Standards in Humanitarian Response* (4th editio). Practical Action Publishing. https://doi.org/ISBN 978-1-908176-00-4

Sphere Project. (2004). The Sphere Project: Humanitarian Charter and Minimum Standards in Humaniratian Response. In *Response* (Second edi, Vol. 1). Oxfam Publishing. https://doi.org/10.3362/9781908176202

Spiegel, P. B., Burkle, F. M., Dey, C. C., & Salama, P. (2001). Developing public health indicators in complex emergency response. *Prehospital and Disaster Medicine*, *16*(4), 281–285. https://doi.org/10.1017/S1049023X00043430

Steets, J., Sagmeister, E., & Ruppert, L. (2016). *Eyes and Ears on the Ground: Monitoring aid in insecure environments*. Global Public Policy Institute. http://www.gppi.net/fileadmin/user_upload/media/pub/2016/SAVE__2016__Monitoring_aid_in_insecure_environments.pdf

Stewart, S. (2014). Design Research. In *The SAGE Encyclopedia of Action Research* (pp. 246–247). SAGE. https://doi.org/10.4135/9781446294406

Suzman, J. (2017). *Affluence Without Abundance: The Disappearing World of the Bushmen*. Bloomsbury.

Tansey, C. M., Pringle, J., Davé, A., Boulanger, R., & Hunt, M. (2018). Earthquakes to Floods: A Scoping Review of Health-related Disaster Research in Low- and Middle-income Countries. *PLoS Currents*. https://doi.org/10.1371/currents.dis.57d98a902a326361d88d54521e68b016

Tatman, R. (2017). Gender and Dialect Bias in YouTube's Automatic Captions. *Proc. of the First Workshop on Ethics in Natural Language Processing*, 53–59. http://www.aclweb.org/anthology/W17-1606

Tatoeba. (2022). *Tatoeba*.

Taylor, B., Henshall, C., Kenyon, S., Litchfield, I., & Greenfield, S. (2018). Can rapid approaches to qualitative analysis deliver timely, valid findings to clinical leaders? A mixed methods study comparing rapid and thematic analysis. *BMJ Open*, *8*(10), 5–7. https://doi.org/10.1136/bmjopen-2017-019993

Taylor, L. (2016). The ethics of big data as a public good: which public? Whose good? *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *374*(2083), 20160126. https://doi.org/10.1098/rsta.2016.0126

Tegmark, M. (2017). *Life 3.0 : being human in the age of artificial intelligence*. Alfred A. Knopf.

Terry, F. (2002). *Condemned to repeat? : the paradox of humanitarian action*. Cornell University Press.

The Engine Room and Oxfam. (2018). *Biometrics in the Humanitarian Sector*. https://www.theengineroom.org/wp-content/uploads/2018/03/Engine-Room-Oxfam-Biometrics-Review.pdf

The Ethics Centre. (2020). *What is Ethics?* The Ethics Centre. https://ethics.org.au/why-were-here/what-is-ethics/?gclid=Cj0KCQiA7aPyBRChARIsAJfWCgJX4UMObXi-8mocGDgBBLPGNH9mc5-cIBBqh-zgFBnX6TBQLpooDoIaAmWlEALw_wcB

The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems. (2019). *Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems* (First edit). https://doi.org/10.1109/MCS.2018.2810458

Thorley, L., & Henrion, E. (2019). *Ethical standards rapid literature review*. DFID. https://assets.publishing.service.gov.uk/media/5d5e3df5ed915d08d612d835/DFID_Ethics_Literature_Review_FINAL_30_05_19.pdf

Thylin, T., & Duarte, M. F. N. (2019). Leveraging blockchain technology in humanitarian settings - opportunities and risks for women and girls. *Gender and Development*, *27*(2), 317. https://doi.org/http://dx.doi.org/10.1080/13552074.2019.1627778

Translators Without Borders. (2018a). *The words between us: How well do enumerators understand the terminology used in humanitarian surveys? A study from northeast Nigeria*. https://translatorswithoutborders.org/wp-content/uploads/2018/12/TWB_Nigeria_EnumeratorComprehension_Nov2018-1.pdf

Translators Without Borders. (2018b, April 19). *Communications Dashboard: Internally Displaced People in Northeast Nigeria*. https://translatorswithoutborders.org/communications-dashboard-internally-displaced-people-in-north-east-nigeria/

Translators Without Borders. (2019). *Language mapping: Putting communication needs on the map*. https://translatorswithoutborders.org/language-mapping/

Travis, D., & Hodgson, P. (2019). *Think Like a UX Researcher: How to Observe Users, Influence Design, and Shape Business Strategy* (1st ed.). CRC Press.

Trials of war criminals before the Nuernberg military tribunals under Control Council law no. 10 (Volume 2). (1949). U.S. Government Printing Office. https://collections.wip.nlm.nih.gov/catalog/nlm:nlmuid-01130400RX2-mvpart

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D. J., Horsley, T., Weeks, L., Hempel, S., Akl, E. A., Chang, C., McGowan, J., Stewart,

L., Hartling, L., Aldcroft, A., Wilson, M. G., Garritty, C., … Straus, S. E. (2018). PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. *Annals of Internal Medicine*, *169*(7), 467. https://doi.org/10.7326/M18-0850

Tricco, A. C., Zarin, W., Rios, P., Pham, B., Straus, S. E., & Langlois, E. V. (2016). Barriers, facilitators, strategies and outcomes to engaging policymakers, healthcare managers and policy analysts in knowledge synthesis: a scoping review protocol. *BMJ Open*, *6*(12), e013929. https://doi.org/10.1136/bmjopen-2016-013929

UN General Assembly. (1948). Universal Declaration of Human Rights. In *General Assembly resolution*. https://www.un.org/en/ga/search/view_doc.asp?symbol=A/RES/217(III)

UN Global Pulse. (2016). *Big Data for Development and Humanitarian Action: Towards Responsible Governance: Global Pulse Privacy Advisory Group Meetings*. http://unglobalpulse.org/sites/default/files/Big_Data_for_Development_and_Humanitarian_Action_Report_Final_0.pdf

UN News Centre. (2015, February 6). *Interview: Senior UN relief official warns no room for complacency until Ebola is eliminated*. https://ebolaresponse.un.org/interview-senior-un-relief-official-warns-no-room-complacency-until-ebola-eliminated

UN OCHA. (2003). *Glossary of Humanitarian Terms*. OCHA. https://www.humanitarianresponse.info/sites/www.humanitarianresponse.info/files/documents/files/ocha glossary.pdf

UNHCR. (2006). *Master Glossary of Terms Rev. 1 United Nations High Commissioner for Refugees*. Geneva: UN High Commissioner for Refugees (UNHCR). http://www.refworld.org/docid/42ce7d444.html

UNHCR. (2016). *Connecting Refugees: How Internet and Mobile Connectivity can Improve Refugee Well-Being and Transform Humanitarian Action*. http://www.unhcr.org/5770d43c4.pdf

UNHCR. (2017a). *Needs Assessment Handbook*. http://needsassessment.unhcr.org/media/downloads/2017-05_NAH_BOOK_NEW.pdf

UNHCR. (2017b). *Standard WASH KAP Questionnaire*. http://wash.unhcr.org/download/wash-kap-survey-module-0-general/

UNHCR. (2018). *Global Trends 2017*. https://www.unhcr.org/5b27be547.pdf

UNHCR. (2021). *Global Trends: Forced Displacement in 2020*. UNHCR. https://www.unhcr.org/60b638e37/unhcr-global-trends-2020

United Nations. (2016). *One humanity: shared responsibility. Report of the Secretary-General for the World Humanitarian Summit*. New York: UN General Assembly. https://documents-dds-ny.un.org/doc/UNDOC/GEN/N16/025/68/PDF/N1602568.pdf?OpenElement

United Nations. (2022). *Report of the UN Secretary-General on the Protection of Civilians in Armed Conflict (S/2022/381)*. United Nations. https://documents-dds-ny.un.org/doc/UNDOC/GEN/N22/321/73/PDF/N2232173.pdf?OpenElement

Vaishnavi, V. K., & Kuechler, W. (2015). Chapter 2: Introduction to Design Science Research in Information and Communication Technology. In *Design science research methods and patterns: Innovating information and communication technology* (2nd ed.). CRC Press. https://doi.org/10.1201/9781420059335

Van der Windt, P., & Humphreys, M. (2016). Crowdseeding in Eastern Congo. *Journal of Conflict Resolution*, *60*(4), 748–781. https://doi.org/10.1177/0022002714553104

Van Mierop, E. S. (2015). Coming clean on neutrality and independence: The need to assess the application of humanitarian principles. *International Review of the Red Cross*, *97*(897–898), 295–318. https://doi.org/10.1017/S181638311500065X

van Wynsberghe, A., & Comes, T. (2020). Drones in humanitarian contexts, robot ethics, and the human–robot interaction. *Ethics and Information Technology*, *22*(1), 43–53. https://doi.org/10.1007/s10676-019-09514-1

Vannini, S., Gomez, R., Lopez, D., Mora, S., Morrison, J. C., Tanner, J., Youkhana, L., Vergara, G., & del Mar Moreno Tafurt, M. (2020). Humanitarian organizations' information practices: Procedures and privacy concerns for serving the undocumented. *Electronic Journal of Information Systems in Developing Countries*, *86*(1), 1–8. https://doi.org/10.1002/isd2.12109

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). *Attention Is All You Need*. https://arxiv.org/pdf/1706.03762.pdf

Vincent, J. (2019). *AI researchers debate the ethics of sharing potentially harmful programs*. The Verge. https://www.theverge.com/2019/2/21/18234500/ai-ethics-debate-researchers-harmful-programs-openai

Vindrola-Padros, C., & Johnson, G. A. (2020). Rapid Techniques in Qualitative Research: A Critical Review of the Literature. *Qualitative Health Research*, *30*(10), 1596–1604. https://doi.org/10.1177/1049732320921835

Wade, N. (2006). Before the dawn: recovering the lost history of our ancestors. In *Choice Reviews Online* (Vol. 44, Issue 02). Penguin Books. https://doi.org/10.5860/choice.44-1018

Walker, P. (2005). Cracking the code: The genesis, use and future of the Code of Conduct. *Disasters*, *29*(4), 323–336. https://doi.org/10.1111/j.0361-3666.2005.00295.x

Walker, P., & Purdin, S. (2004). Birthing sphere. *Disasters*, *28*(2), 100–111. https://doi.org/10.1111/j.0361-3666.2004.00246.x

Wang, N., Christen, M., & Hunt, M. (2021). Ethical Considerations Associated with "Humanitarian Drones": A Scoping Literature Review. *Science and Engineering Ethics*, *27*(4), 1–21. https://doi.org/10.1007/s11948-021-00327-4

Watkins, D. C. (2017). Rapid and Rigorous Qualitative Data Analysis: The "RADaR" Technique for Applied Research. *International Journal of Qualitative Methods*, *16*(1), 1–9. https://doi.org/10.1177/1609406917712131

Watson, H., & Rodrigues, R. (2018). Bringing Privacy into the Fold: Considerations for the Use of Social Media in Crisis Management. *Journal of Contingencies & Crisis Management*, *26*(1), 89–98. http://10.0.4.87/1468-5973.12150

Weiss, R. J., Chorowski, J., Jaitly, N., Wu, Y., & Chen, Z. (2017). Sequence-to-sequence models can directly translate foreign speech. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, *2017-Augus*, 2625–2629. https://doi.org/10.21437/Interspeech.2017-503

WFP. (2022). *Innovation Accelerator*. https://innovation.wfp.org/

WHO. (2007). *Risk reduction and emergency preparedness*. World Health Organization. https://apps.who.int/iris/bitstream/handle/10665/43736/9789241595896_eng.pdf

WHS. (2016). *The Grand Bargain – A Shared Commitment to Better Serve People in Need*. Istanbul: World Humanitarian Summit. http://reliefweb.int/sites/reliefweb.int/files/resources/Grand_Bargain_final_22_May_FINAL-2.pdf

Wood, G. (2018, April). The Refugee Detectives. *The Atlantic*. https://www.theatlantic.com/magazine/archive/2018/04/the-refugee-detectives/554090/

World Bank. (2020). *World Bank Country and Lending Groups*. The World Bank. https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups

World Medical Association. (2013). World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *JAMA*, *310*(20), 2191–2194. https://doi.org/10.1001/jama.2013.281053

XLSForm.org. (n.d.). *XLSForm.org*.

Yeophantong, P. (2014). Understanding Humanitarian Action in East and Southeast Asia: A Historical Perspective (London: ODI, 2014). In *HPG Working Paper* (Issue February). ODI/HPG. http://www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/8807.pdf