Mapping and Comparing Responsible Data Approaches

JUNE 2016

Jos Berens *Center for Innovation*
Ulrich Mans *Center for Innovation*
Stefaan Verhulst *The GovLab*

GOVLAB

CENTRE FOR INNOVATION
Leiden University

TABLE OF CONTENTS

A. The need for mapping and comparing responsible data approaches 3

B. Methodology 5

B.1. Included Data Responsibility Approaches 5

B.3. Generic insights gained from the peer review. 7

C. Comparative analysis and findings. 10

C.1. Scope of the Policy 10

C.2. Value-Proposition 12

C.3. Data 14

C.4. Risk-Assessment 16

C.5 Value -Chain 22

C.6. Principles and Legal Foundation 23

C.7. Tools and Practices 24

C.8. Accountability and Design 25

Appendix 1. Repository of Documents Reviewed 31

Appendix 2. Comparative Table. 35

Mapping and Comparing Responsible Data Approaches[1]

JOS BERENS, ULRICH MANS AND STEFAAN VERHULST[2]

A. The need for mapping and comparing responsible data approaches

We are witnessing a growing awareness of the potential offered by data in the humanitarian space. Data, and more generally information, can offer insights into fast-moving situations, helping humanitarian organizations, policymakers and others identify emerging crises, track their spread or evolution, and respond in more targeted and effective ways. At the same time, the exponential increase in the amount of available data—and in the sources of data—adds a new level of complexity. As much as data offers new opportunities, it also poses new risks—risks that are magnified given the vulnerability of affected stakeholders. It is therefore essential to ensure that any use of data in humanitarian contexts is governed through a balanced and well-articulated set of data policies and guidelines. In order to realize the potential of data while minimizing its harms, we need a framework for data responsibility.

Such a framework is key to ensuring a fair and equitable approach to the use of data. Perhaps most importantly, it is essential to ensuring equity in the on-the-ground impact of policy decisions and actions that rely on data. Inadequate or irresponsible handling of data does not only pose the risk of ineffectiveness; it can also negatively impact "data subjects," threatening further harms to the very populations that data is supposed to serve. Therefore, in order to create adequate level of trust and ensure the effectiveness

1 The comparative analysis was commissioned by United Nations Office for the Coordination of Humanitarian Affairs (UN-OCHA) to inform the development of data policies at UN-OCHA and engage the broader humanitarian community on data responsibility. The analysis and subsequent opinions and recommendations expressed in the paper are the authors' own and do not reflect the view of UN-OCHA.

2 We are grateful to Melissa Amoros-Lark and Nicolas Castellon (Student Fellows at Leiden University's Centre for Innovation) for their research assistance.

of data-driven innovations across the humanitarian sector, data policies, guidelines and implementation safeguards need to be developed and rigorously tested. An ability to measure (and, if necessary, adjust) the outcome of actions is also essential: it must be clear how to close loopholes and correct unsuccessful practices in order to hold governing agencies and others accountable.

The need for a responsible data governance approach has now been recognized by a wide variety of organizations, both within and outside the humanitarian space.[3] In addition, there are various technology-oriented organizations that operate in related fields and that are also working on in developing their responsible data approach, humanitarian organizations should be guided both by the themes found in other data policies and guidelines, and by other documents that determine how they should conduct their work (such as the humanitarian principles) that have started considering responsible data use or that seek to revise existing policies to adapt to the increasingly fast-paced data landscape.[4]

The goal of this paper is to examine some of these existing approaches and, based on a comparative analysis, to identify best practices and innovative approaches to governing data in humanitarian contexts. To that end, Leiden University's Centre for Innovation and the Governance Laboratory at New York University (The GovLab) have together undertaken a mapping exercise of 17 existing responsible data approaches. The list of organizations studied and links to their data responsibility approaches can be found in Appendix 1: it includes 7 UN agencies, 7 International Organizations, 2 government agencies and 1 research institute.[5] The results of our comparative exercise are presented in Section C, which presents an analysis organized along eight key themes. Section D includes six takeaways or best principles. Taken together, these can serve as a toolkit for any organizations—particularly those operating in the humanitarian space—seeking to use data more responsibly and effectively.

This work is part of a larger collaboration between the GovLab, C4i, Data-Pop Alliance, UN Global Pulse and Data & Society Research Institute - the International Data Responsibility Group. The IDRG is a global network of experts and organizations working on the principles and standards that are required for guiding the Data Revolution in the context of humanitarian action, sustainable development and peace & justice. Its members seek to build an authoritative knowledge platform that enables responsible experimentation on the release, processing and use of data and minimising risks. The IDRG is designed as a networked platform, with a coordinating secretariat in The Hague, The Netherlands.

3 See for example the Resolution adopted during the 37th International Conference of Data Protection and Privacy Commissioners, to be found at: *https://icdppc.org/wp-content/uploads/2015/02/Resolution-on-Privacy-and-International-Humanitarian-Action.pdf*

4 Examples include the United Nations Global Pulse team (Privacy Impact Assessment), World Food Programme and International Organization for Migration.

5 Repository includes documents by: EU Regulation, GSM Association, IOM, LIRNEasia, Médecins Sans Frontières, Principles for Digital Development, Oxfam, UN Global Pulse, UNHCR, UNICEF, the Humanitarian Principles, HDX's policy, UN Office for Outer Space Affairs, United Nations Population Fund, USAID, ICRC and the White House Precision Medicine Initiative.

B. Methodology

As noted, this paper contains a comparative analysis of 17 responsible data approaches–including data policies, principles and regulations (see summarized list in B.1). These approaches were selected to reflect developments in data responsibility in international organizations of various sizes and with different missions. They add up to something of a map of existing efforts that can help lay the foundations for our efforts to develop a set of best practices that can guide other organizations in their efforts to use data more responsibly and effectively.

The 17 approaches selected are not meant to provide a comprehensive overview of all existing approaches, yet were selected because each represents a particular model, approach or set of principles that if taken together could enable the development of a meaningful data policy. In addition, several approaches are of immediate relevance to the humanitarian field while others are considered landmark documents that provide a foundation to build upon.

In analyzing the respective approaches, we use the template outlined in B.2. This template, which has been jointly developed by The GovLab and Leiden includes the key components of a responsible data governance framework, and provides a full matrix of analysis to identify best practices.

B.1. INCLUDED DATA RESPONSIBILITY APPROACHES

Médecins Sans Frontières	*Data Sharing Policy*	*2013*
Oxfam	*Responsible Program Data Policy*	*2015*
UN Population Fund	*Information Disclosure Policy*	*n.d.*
UN-OCHA	*Humanitarian Data Exchange Terms of Service*	*n.d.*
UNHCR	*Policy on the protection of Personal Data of Persons of Concern to UNHCR*	*2015*
LIRNEasia	*Draft Guidelines for Third-Party Use of Big Data Generated by Mobile Network Operators*	*2014*
GSMA	*Guidelines on the protection of privacy in the use of mobile phone data for responding to the Ebola outbreak*	*2014*
White House	*Precision Medicine Initiative; Privacy and Trust Principles*	*2015*
UN Global Pulse	*Privacy and Data Protection Principles*	*2015*

UN Office for Outer Space Affairs	*International Charter for Space & Major Disasters*	*2000*
UN-OCHA	*Humanitarian Principles*	*1991, 2004*
Digital Impact Alliance	*Principles for Digital Development*	*n.d.*
European Union	*Draft General Data Protection Regulation*	*2012*
International Organization for Migration	*Data Protection Policy*	*2010*
UNICEF	*Information Disclosure Policy*	*2011*
USAID ADS Chapter 508	*Privacy Program*	*2014*
International Committee of the Red Cross	*Rules on Personal Data Protection*	*2016*

B.2. TEMPLATE OF ANALYSIS

Based on previous research undertaken inside and outside this field, and based on a preliminary assessment of both the selected data responsibility approaches and additional ones, GovLab and Centre for Innovation distilled the key elements of a well-balanced and comprehensive data responsibility approach. Based on these essentials, GovLab and Leiden developed the following template for analysis:

SCOPE OF THE POLICY
Does the policy clarify its goals and objectives? (For example, does it clearly state objectives such as 'data protection', 'data subject safety', 'organisational integrity', 'infrastructural security' and others)
Does the policy describe the context in which data is collected, used and shared (clarifying, for example, data types, origin and processing method)?

VALUE-PROPOSITION
Does the policy indicate the expected value or benefit of the protection, use and analysis of the data-set(s)? And if so, how?
Does the policy include suggestions on how the impact and value of the data will be measured? And if so, how?

DATA
Does the policy describe the (technical specifications – i.e. the type of files, the size, and other characteristics of) data-set(s) it oversees and if so how?
Does the policy include an audit or process to determine what data is necessary for the purpose and anticipated value?

RISK-ASSESSMENT
Does the policy indicate or establish a process to determine the risk that the analysis and/or use of the data may generate, either to the organization, its beneficiaries or others?
Does the policy describe or list risk mitigation strategies to mitigate or respond to the risk? If so, which ones are listed?

VALUE-CHAIN
Does the policy describe the value-chain of data and the benefit/risks at each stage?
How is the data processed?

PRINCIPLES and LEGAL FOUNDATION
Are the principles (and or ethical norms) explained that guide the policy?
Is the legal basis upon which the policies and principles build explained? And what is that basis?

TOOLS AND PRACTICES
What tools and practices are specified to implement the principles and policies?
What decisions do those tools and practices inform?

ACCOUNTABILITY and DESIGN
Does the policy explain how it was created?
Is it based upon either participatory or user design principles?
What monitoring and evaluation mechanisms are implemented by this data strategy?
Are there dispute resolution mechanisms?
Does the policy explain any roles and functions that are tasked with the implementation of the policy?

The selected data responsibility approaches were analyzed according to the above template, and were checked for their inclusion of the themes. A quantitative comparison revealed the prevalence of themes in the selection of approaches. Subsequently, qualitative analysis was directed particularly towards unexpected outliers in the quantitative comparison.

B.3. GENERIC INSIGHTS GAINED FROM THE PEER REVIEW

The authors of this paper are thankful for the insights and feedback from our review group. We received and integrated comments from the ICRC, IOM, LIRNEasia, UN Global Pulse, UNHCR, UNICEF and USAID. All reviewers acknowledged the need to develop a data governance framework, and provided several comments and insights on our effort that have proven invaluable in taking this project forward. We integrated all their review, yet some reviewers provided some important top-level reflections, including:

> *"The document that has been distributed for comment seeks to impose the European approach on emerging 'big data' practices, ignoring the alternative approaches and indeed stripping away essential elements of the context-specific guidelines adopted in several of the documents covered by the interim findings (...)"*

"We are in the process of discovering the potential of big data for humanitarian purposes. Overly prescriptive and rigid frameworks derived from entirely different circumstances (use and abuses of information pertaining to creditworthiness) to an inchoate field of investigation has the potential to stifle discoveries. To minimize harm, it is advisable to be minimalistic in devising regulatory schemes."

"There don't seem to be any reflections within existing policies on the particular challenges associated with merging of databases or interoperability. The assumption seems to be that most data sources are singular but the merging of datasets provides both the greatest challenges to privacy as well as the greatest opportunities."

"(...) the term 'responsible data approaches' is a very generic term which hides a lot of variation and nuances. In fact governance mechanisms within a specific institution may differ hugely depending upon whether the data is big/open; actively or passively collected; generated in an online or offline environment as well as the local context where the data was collected e.g 'right to be forgotten principle' is not universal. It might be useful for the authors to reflect on the need for greater disaggregation and categorization within a broad data governance mechanism."

"It would be interesting to reflect on how many of the policies focus more on data collection and sharing rather than longer term management, updating and storage. This includes the need for ongoing consent or an acknowledgement that this may be required which I think is generally poorly reflected. I think there is also a general absence of references in policies to virtual data storage."

'Responsible', 'data' and 'approach' are all very broad terms, causing the risk of a too generic and high-level comparison.

"This is not only about ethics. There is (also) law. (...) based on the 1990 UN GA Guidelines and Human Rights Law which even IGOs cannot entirely ignore where their activities have an impact on individuals (...), there is a 'residual' body of law that they should respect when dealing with personal data."

"On Innovative approaches and the consideration of behavioral or economics approaches to inform particular choices or behaviors. This is definitely interesting but should probably be notes as complementary to regulatory approaches. May need a combination of 'carrots' and 'sticks' when systemic institutional change is required."

"(..) an issue we have heard repeatedly from development actors is lack of clarity about where the policy sits within the organization in terms who is ultimately accountable for implementation of the policy and to revisit the policy if needed."

"Change in organizational mindset is also a big challenge when developing such policies … [Need] to be mindful of this and ensure that organizational needs are clear upfront and stakeholders see the benefit of such a policy to their work, particularly, in humanitarian contexts where it could be seen as administratively burdensome in conflict situations."

International organisations developing their data strategy cannot choose between a high-level approach or implementation documents if they are not subject to domestic law, and will need both a policy and mechanisms for implementation.

SOME RESPONSES TO THESE COMMENTS

This work intentionally took a broad approach in including relevant documents for analysis. We felt it important to be able to provide a template of analysis that could be made operational for humanitarian organizations seeking to draft a data policy. Seeking out recurring *themes* instead of similar clauses or phrases allowed the team to include principles, guidelines and policies. Keeping in mind the rapid development of the field of humanitarian data use, these meta-results leave ample room to determine the most feasible, workable and practical modality for implementation. Further, such a broad approach provides a sense not only of the themes to be covered, but also the different types of modalities to address them.

Second, we didn't aim to be fully comprehensive of all the data policies that exist yet we curated approaches that could provide lessons, practices and language relevant for humanitarian organizations. Based on this curated approach, the authors aimed to capture an initial base of lessons learned, which was reviewed and improved through a participatory process. In the review process, several additional documents were suggested for inclusion, and the authors will take these resources into account for any future follow-up to this work.

Finally, this work is not a legal analysis, although the applicable legal context in which a policy is created and operationalized should always be taken into account. Given that many humanitarian organizations work across borders, they may need to comply with varying local legal regimes or different local implementations of international data protection law.

C. Comparative analysis and findings

In what follows, we compare the 17 documents along eight themes. These eight themes constitute what we believe to be the key elements of a responsible data use framework. They are designed to be broad enough to be broadly applicable, yet specific enough to be operational and actually usable.

C.1. SCOPE OF THE POLICY

Determining the scope of a data policy–i.e. to which data, which use and which actors the policy applies– is important for data users to determine whether the policy applies to them, and how; and for partner organisations in assessing how their data may be handled if they share it. It also provides an opportunity to share the rationales, goals and priorities behind the policy.

C.1.1. CLARIFICATION OF GOALS AND OBJECTIVES

All of the reviewed data policies include a section that explains the overall goal of the document. There is a broad variety in the way goals are stated, ranging from very concrete user cases to generic principles.

- The Global System for Mobile Communication Association (GSMA) guidelines, for example, were developed for mobile data sharing efforts aimed at fighting the Ebola epidemic,
- While MSF emphasizes the organization's wish to ensure the highest standards in monitoring and documentation.
- The International Organization for Migration (IOM) data protection manual features the most detailed list of goals ("key objectives"), including
  - data-specific aspects (among others integrity, confidentiality, data protection) as well as;
  - organizational aspects (among others institutional safeguards, enhance understanding).
  - Other data policies showed forward-looking benefits such as preparing for the rising use of data (e.g. Oxfam's data policy states it is preparing the organization for the future) and to build trust in the online environment as to aid the digital economy (e.g. EU General Protection Data Regulation).

In general, we identified two types of goals:

- Those related to the impact on the use of data that the policy aims to facilitate, such as the use of data for monitoring and documenting interventions to improve services (see also value proposition below),[6] and;

6 See for example the Medicins Sans Frontieres Data Sharing Policy, in which MSF states that they "place a high value on monitoring and documenting MSF medical interventions in order to continually improve the quality of care delivered."

- Those goals that relate to the *responsible* use of data as a means of protecting data subjects and their rights,[7] preventing liability, risk and harm.[8] Clearly stating the intent of both data use in the first place, accompanied by a description of the aims of the policy document, significantly aids in the interpretation of the provisions that follow.
  IOM's policy takes account of these two goals and encapsulates it into its definition of data protection: "Data protection is the systematic application of a set of institutional, technical and physical safeguards that preserve the right to privacy with the respect to the collection, storage, use, and disclosure of personal data."

C.1.2. DESCRIPTION OF CONTEXT

More than half of data policies analyzed provided detailed descriptions of the context in which their data is collected and used. The descriptions vary depending on the mandate of the institution as it is apparent that organizations with a technical mandate are more exact with their definitions. There are three categories of documents when it comes to referencing data types (see section 3.1. for technical specifications):

- One category states a specific data type (for example, data on refugees in the case of the UN Office of the High Commissioner for Refugees (UNHCR), humanitarian data in the case of UN-OCHA's Humanitarian Data Exchange (HDX) and mobile data in the case of the GSMA).
- Other institutions cast wider definitions that address the field the data concerns, such as Precision Medicine Initiative stating it concerns medical data, and IOM stating it deals with all types of personal data relating to their beneficiaries. This second category often refers to any data that is relevant in the given context (e.g. "all data used by EU citizens" in the European Data Protection Regulation (EDPR) document and "all information in the possession of the UNFPA [United Nations Populations Fund, ed.]").
- A third category does not mention the type of data that is covered (e.g. *digitalprinciples.org* and Oxfam). In the latter category, one can also include the International Charter: Space and Disasters, which includes a broad definition without using the term *data*: "critical information for the anticipation and management of potential crises".

7 See for example IOM, which states that "The collection and processing of personal data are necessary components of IOM's commitment to facilitate migration movements, understand migration challenges, and respect the dignity and well-being of migrants. IOM's data protection strategy seeks to protect the interests of IOM's beneficiaries, as well as the Organization [...] [i]t recognizes both the rights of individuals to protect their personal data and the need of IOM to collect, use and disclose personal data in the course of fulfilling its migration mandate" – see IOM Data Policy, p.9 and 13.
8 See for example UNHCR, which states that "[i]ts purpose is to ensure that UNHCR processes personal data in a way that is consistent with the 1990 United Nations General Assembly's Guidelines for the Regulation of Computerized Personal Data Files and other international instruments concerning the protection of personal data and individuals' privacy."

C.2. VALUE-PROPOSITION

A description of the purpose and the anticipated benefits of using data is essential for both data subjects, data users, partner organisations and the general public, to understand the benefit/risk balance that an organisation aims to strike through its data policy. When accompanied with concrete metrics and indicators, an expression of the value one hopes to create also allows for subsequent impact assessment–enabling more evidence based policy making.

C.2.1. VALUE PROPOSITION DESCRIPTION

Over half of the data policies reviewed indicate the expected value or benefit of the use and analysis of the data. Benefits were listed in a wide variety of ways such as for:

- improving research (e.g. MSF: *"MSF's large repository of Research data together with routinely collected data can potentially be of value to researchers working in public health."*);
- creating public trust in the work of the organisation through information sharing (UNICEF);
- modelling and planning efforts (GSMA Guidelines for responding to the Ebola outbreak);
- The ICRC lists the following concrete purposes in article 3 of its principles section:
  - restoring family links
  - protecting individuals in detention
  - protecting the civilian population
  - building respect for IHL – including through training and capacity building;
  - providing medical assistance
  - forensic activities
  - weapon decontamination
  - ensuring economic security
  - protecting water and sanitation system
  - preventive and curative health care.
- LIRNEasia remains more general in its value proposition statement, noting in the introduction that "Big Data has an immense, and at this point unique, potential to bring forth a qualitative transformation of urban design including resilience, improve transportation and government-service delivery, and enhance management of the economy, among others."

A general statement of the value proposition of data use is included in the Precision Medicine Initiative Principles, i.e. to enable "a new era of clinical care" without further detail.[9] The IOM Policy contains a full section on the need for a clear value proposition for data use.[10] This includes a mechanism to determine whether the value proposition of additional (unforeseen) use of the data is compatible with the original purpose of use,[11] and 'Compatible Research' on the data, through i.e. 'data matching': *"(...) the electronic comparison of two or more sets of personal data that have been collected for different specified purposes."*[12] Accordingly, assessing the value of the data is also linked to retention and accuracy of the data which may have an impact on service delivery.

C.2.2. DATA IMPACT MEASUREMENT

Our analysis indicates that there is a clear lack of inclusion of value and impact indicators in data policies. Hardly any data policy included suggestions on how the impact and value of the data would be measured, and what indicators may be used to determine success. The "Principles for Digital Development", developed by a collective of international NGOs, UN agencies and donors, mentions the need to design projects so that impact can be measured with discrete milestones to focus on outcomes, yet most fail to do so. The IOM policy states that *"due to the multifaceted nature of IOM's activities, data protection issues need to be considered at all stages, from project development and implementation to evaluation and reporting"*. Yet indicators are not included, instead, this is presumably left to the indicators in the different projects covering the various IOM activities.

Making value and impact indicators explicit allows data users to determine how their data benefits the organisation in a given context. It also allows for more rigorous project evaluation, which will inform the development of subsequent data-driven engagements.

When developing a data governance framework, including indicators is a useful instrument to help determine whether data use is justifiable based on the intended deliverables, especially when sensitive data is accessed in crisis situations. Fast decision making requires clear and concise prescriptions to determine whether to move forward or to halt a given project.

9 In the Precision Medicine Initiative Privacy and Trust Principles: "Precision medicine is enabling a new era of clinical care through research, technology, and policies that empower patients, researchers, and providers to work together toward development of individualized care."

10 IOM Data Protection Policy, from p. 25.

11 IOM Data Protection Policy, p. 28, 'Compatible Secondary Purposes'.

12 IOM Data Protection Policy, p. 29, 'Compatible Research'.

C.3. DATA

Describing and defining the data handled by the respective organization enables subsequent risk assessment, and also forces the entity and data users to determine and justify the data they access and use.

C.3.1. DATA TYPES

Most of the documents analyzed contain a description of the type of data to which they are applicable. Depending on the specificity of the document–ranging from highly specific (GSMA Guidelines focused on Call Detail Record use in the Ebola response), to general (e.g. the Precision Medicine Initiative Privacy and Trust Principles)–such descriptions equally varied from technically detailed to high-level and broad outlines of the types of data governed by the document. The below lists the descriptions found.

- The GSMA data policy divides their data in terms of regular "Call Details Records (CDR)" and their Anonymized CDR and the Mobile Network Operator (MNO) goes on to specifying Visitor Location Register (VLR) data, and GPS data.
- UNICEF includes a broader description of 'information' in their Information Disclosure Policy (For purposes of this Policy, 'Information' means any produced content, whatever its medium (paper, electronic or sound, visual or audiovisual recording), concerning a matter relating to the policies, activities and decisions of UNICEF.)
- The International Charter for Space and Major Disasters (ICSMD) specifically defines their "Space Data" as raw data gathered by a space system controlled by a party that can transmit it to a ground receiving station. It also points out the distinction between data and information, defining the latter as: "data that have been corrected and processed by the parties using an analysis program, in preparation for use in crisis management by one or more associated bodies in aid of the beneficiaries; it forms the basis for the extraction of specific products for use on location;".
- Médecins Sans Frontières focuses more on the purpose of the data, stating that 'dataset' means any single dataset or set of Human Samples with associated data, included in the Collection. An MSF Dataset may have been compiled:
  - for a specific, focused Research, or;
  - for health service provision or planning
- USAID contains a full description and background of Personally Identifiable Information (PII) in article 508.3.1. 'PII'.
- For the White House Precision Medicine Initiative, the Privacy and Trust Principles are rendered applicable to all data used by the program, leaving further details on what this data exactly is, to " (...) future PMI activities regarding: governance; transparency; participant empowerment; respect for participant preferences; data sharing, access, and use; and data quality and integrity."

- LIRNEasia's draft guidelines contain a helpful description of the meaning of 'big data' and 'Mobile Network Big Data (MNBD)' in the introduction, the latter consisting of CDR's of voice calls and sms, VLR Data, reload data applicable to prepaid customers, GPS data, and big data generated on mobile devices by the use of certain online services.
- UNFPA describes both what they consider 'information'–"any produced content, whatever its medium (paper, electronic or sound, visual or audiovisual recording), concerning a matter relating to the policies, activities and decisions of UNFPA."–and the type of content that is generally processed–see 8.1 and 8.2.
- The EU Regulation focuses on the data subject, and applies to all data concerning EU citizens.
- The IOM Policy applies to all types of personal data relating to IOM beneficiaries and mentions a sensitivity assessment that needs to be conducted prior to data collection, in order to identify necessary safeguards, access controls and security measures to be applied throughout the life cycle of the data processing.[13]
- The ICRC Rules do not contain a technical description of data types, but do contain a description of what is meant by genetic data, health data and personal data, in the annex.
- Oxfam notes that for the purposes of their policy, "data" is considered to be the "physical representation of information in a manner suitable for communication, interpretation, or processing by human beings or by automatic means. Data may be numerical, descriptive, audio, or visual."

When handling various types of data, with new types often being added on the go, too strict of a description of the data to which the policy applies may make the policy obsolete quickly. At the same time, it is key to differentiate between different categories of data to ensure an appropriate balance of risk and utility. Besides a categorisation of data types to indicate the governance regime applicable, pointing out specific types of data that are either very common (such as satellite data), or that pose particular risks (such as Personally Identifiable Information (PII)) is an effective way of institutionalising both accuracy and flexibility.

C.3.2. PURPOSE AND ANTICIPATED VALUE AUDIT

Introducing a data audit system that determines what data is necessary to achieve the expressed goals and/or solve the problem will enable a more rigorous approach to data–preventing collection, sharing and use of data that may not be necessary but add risks instead.

About a third of data policies analyzed included an audit or process to determine what data was necessary for the purpose and anticipated value. Data policies presented two ways of assessing the data:

- through the stipulated guidelines–which include reference to levels of anonymization necessary for personally identifiable information (PII) to be shared with a wider public, as is the case with UN-

13 IOM Data Protection Policy, p.15.

OCHA HDX. IOM's policy includes data coding, pseudonymization, and anonymization as methods to remove identifiable information to preserve the confidentiality of PII and maintain the anonymity of data subjects while at the same time allowing for continued use and value if needed for further use internally or for sharing externally.

- through a Data Custodian or Controller function or an auditing body–placing the responsibility onto a person or group of people in the organization, as is the case in UNHCR's "Data Controller", and MSF's "Data Custodian" where their responsibilities are to grant access to the data, oversee its management, determine the purpose, oversee the processing of the data and have final accountability for the security and integrity of the data.

C.4. RISK-ASSESSMENT

Incorporating (automated) risk assessment methods in data use is vital to ensuring the proportionality of using data to achieve an intended benefit, on a more practical and operational level than a policy can achieve.

C.4.1. RISK ASSESSMENT PROCESSES

About one third of data policies indicated or established a process to determine the risk the data may generate. There are three categories of documents:

- those that state the importance of considering risk–Most of the policies analysed contained a general statement on risk, which is valuable in itself, but clearly insufficient to prevent risk from materializing. The Precision Medicine Initiative states, for instance, that risks and potential benefits of data subjects should be considered;
- those referring to specific risk assessment tools, without including these tools in the policy. The ICRC for example refers to the Data Protection Impact Assessments (DPIA), which are contained in separate documents.
- those that include concrete risk assessment tools–organizations with integrated risk assessments construct check-lists to measure the risks of using the data. IOM, for instance, has developed a "sensitivity" scale ranging from "High", "Moderate" and "Low" for their data controller to assess the data collection and processing along with set of key considerations such as potential for harm and discrimination. Further, the IOM Policy recommends conducting a risk assessment, as a value judgement, prior to data collection and throughout the data processing process to ensure that adequate data protection safeguards are in place at all times.[14]

14 IOM Data Protection Manual, p.141-147.

Though some data policies may not contain a reference to risk assessment, they contain descriptions of privacy issues, such as surveillance, aggregation, identification, insecurity, exclusion, breaches, and disclosure, as stated for example by the MNO data policy. The Médecins Sans Frontières policy includes several examples of types of sensitive datasets on page 8:

> *"(i) Any data from which an implication of criminal conduct could be drawn and/ or that can put Data Subjects at serious risk (including death); this includes data on violence-related medical activities particularly but not exclusively in contexts of conflicts: (1) any data related to violence – bullet wounds; and (2) any data related to sexual violence;*
>
> *(ii) Data collected from MSF activities in prisons or any situation that can be assimilated to detention or deprivation of liberty (including certain refugees or displaced settings);*
>
> *(iii) Certain data variables such as those that could indirectly imply, truly or not, racial or ethnic origin; political or religious opinions (for example from the origin or the location of the Data Subject);*
>
> *(iv) Data related to sicknesses with an obligation to abide to treatment.*

Data potentially considered as Sensitive Data by MSF (non exhaustive):

> *(i) Data that can put Data Subjects at risk of stigma, discrimination and even criminal sanction (including in certain countries or groups of population, HIV and TB data);*
>
> *(ii) Data on sicknesses or epidemic outbreaks."*

USAID has a privacy impact assessment template attached to its policy as addendum.[15] Not yet published and therefore not included in this analysis is the Privacy Impact Assessment that is currently being developed by UN Global Pulse, which could provide relevant insights for humanitarian organizations seeking to use data. Not included in this analysis for the same reason is the forthcoming Data Security Impact Assessment by ICRC, which could provide further insights.

15 See: *https://www.usaid.gov/sites/default/files/documents/1868/508mac.pdf*

Especially when dealing with sensitive data on crisis-affected communities, a strong risk assessment framework is essential to prevent disproportionate data use. The ICRC DPIA may provide a useful example.[16] ICRC did not include a definition of 'sensitive data' in their policy, since this definition is highly contextual. The DPIA, the use of which is required by article 17 of the ICRC rules, sets up a comprehensive framework for the assessment of risk on the following themes:

- Purpose specification;
- Data limitation;
- Right to information;
- Legal basis for data processing / transfer;
- Right to access / rectification / deletion;
- Information quality and accuracy;
- Appropriate security measures;
- Data sharing, disclosure/publication and/or transfer;
- Data retention;
- Risks to individuals;
- Accountability / oversight mechanism.

16 The ICRC Data Protection Impact Assessment template is available here: [https://www.icrc.org/en/download/file/18149/dpia-template.pdf].

C.4.2. RISK MITIGATION AND RESPONSE

Risk mitigation strategies are referred to in almost all the documents analyzed, and to some extent many of the provisions included in these documents attempt to mitigate risk by prohibiting a variety of data uses, by preventing access and security violations, and by promoting appropriate data management. There are numerous ways in which organizations can deal with risk mitigation. UN-OCHA HDX has an extensive risk mitigation strategy that includes aggregation or de-identification, setting strict data sharing policies, obtaining informed consent, and avoiding the duplication of the data collection. IOM, too, devotes substantial attention to the assessment of benefits and risks on a continual basis (see below under 5.1), listing the following risk reducing measures:

- **Elimination:** removing the risk is the safest and best way to reduce the risk.
- **Substitution:** substituting the hazard with something less risky is the best alternative if elimination is impossible.
- **Containment:** using strict supervisory controls can help minimize the likelihood of harm occurring.
- **Reducing exposure:** taking extra precautions against unnecessary exposure can reduce the likelihood of harm occurring.
- **Training:** raising awareness at collection sites can assist with identifying and managing risks.
- **Monitoring:** continuous monitoring can help identify appropriate safeguards to minimize the risk.[17]

Other organizations express the need and intention to implement risk mitigation strategies, such as UNHCR's data policy which expresses the need to implement appropriate organizational and technical measures to meet the requirements of their policy. These include the implementation of principles privacy by design and by default.[18] In terms of data collection, Oxfam states that its agents do not collect non-essential data as to not put participants at risk and will dedicate more efforts into mitigating risks for participants from vulnerable populations.

Data policies with organizational measures set up extensive protocols that consist of procedures, trainings, and assessments. UNHCR's "Data Security Measure" describes standard operating procedures, staff trainings in data protection and security, data protection impact assessments, and also addresses the physical security of the equipment. IOM has separate chapters on consent, confidentiality, data security measures, conditions for transfer to third parties, and refers to data protection focal points to assist in monitoring and training as well as data protection audits.

It becomes clear from the analysis that risk mitigation takes place both at the organizational and technical levels. As far as technology is concerned, attention is often paid to the security of the equipment and the

17 See IOM Data Protection Manual 2010.

18 See the ICRC rules article 16 under 1: "While designing a database and drafting procedures for collecting Personal Data, all these rules must be taken into account and incorporated to the greatest extent possible; this is known as "data protection by design and by default."

storage device of the data. In organizational terms, the risk mitigation strategy has been described to be an all-hazards scenario of the potential risks with corresponding measures to prevent them.

The following taxonomy of risk mitigation tools and methods can be discerned. Which strategy is the appropriate one to apply will depend on the risks assessed in the given circumstance:

TECHNICAL PROCEDURES

- **Physical Security (MSF, USAID & UN Global Pulse):**
  - Maintaining physical security of premises;
  - Portable equipment;
  - Individual case files and records;
- **Access Control (UNHCR, UN Global Pulse, IOM & UN-OCHA):**
  - Password management applications;
  - User control;
  - Storage control;
  - Communication and transport control;
- **Classification of Data (UNCHR, UN Global Pulse & IOM):**
  - Sensitivity assessments to determine the sensitivity of data based on factors depending on the specific context in which an organisation operates;
  - A data classification system, ranking the potential to harm or discriminate against the data subject or third-individuals as ‘low’, ‘moderate’ or ‘high’;
- **Security and Encryption (UNCHR, UN Global Pulse, IOM, GSMA):**
  - Anonymize phone numbers of subscribers receiving calls or text messages by assigning anonymous code before analysis using the secure SHA-3 algorithm;
  - Encrypt web-based services using SSL;
  - Offer two-factor authentication for log-ins;
  - Provide adequate firewall and antivirus protection to devices and networks the data is used on;

ORGANIZATIONAL PROCEDURES

- Standard Operating Procedures that include (UNHCR, UN Global Pulse, IOM, UN-OCHA, GSMA, EU, USAID & ICRC):
  - Conducting Impact Assessments;
  - Direct contact point with data controller for inquiries or requests to remove data;
- Legality of Data (UN Global Pulse, UN-OCHA, Oxfam, IOM, EU, USAID, ICRC, & MSF):
  - Only store, access, analyze or otherwise use data in accordance to the purpose for which they were properly and lawfully obtained;[19]
  - Obtain data with informed consent from individuals or communities.[20] When consent cannot be given, provide appropriate safeguards including a privacy and impact assessment;[21]
- Protect User Privacy (MSF, UNHCR, Principles for Digital Development, UN Global Pulse, Precision Medicine Initiative, LIRNEasia, UN-OCHA, UNFPA, Oxfam, GSMA, EU, IOM, & USAID):
  - Protect against the following privacy issues: Surveillance, (inappropriate) aggregation, exclusion, confidentiality breach, increased accessibility, identification of individuals or groups, and secondary usage of the data;
  - Measures to avoid duplication of data collection in order to avoid unnecessary burden to data subjects;
- Data Sharing Protocols (MSF & Precision Medicine Initiative):
  - Share Public Data under a user-selected creative commons license or as a public domain;
  - Private data should only be accessible to member of the organization;
  - No analysis should single out identifiable individuals / groups / communities and no attempt should linked the data to individuals;
  - Only final analysis of data should be made available to aid organizations, government agencies or research institutes;
- Governance mechanisms (Precision Medicine Initiative, UNICEF, UN Global Pulse, UNFPA & USAID):
  - Ensure accountability;
  - Responsible data management;
  - protection against any intentional or unintentional unauthorized access, use, disclosure, or re-identification of data.

19 The ICRC rules make the rationale for including a description of these purposes explicit, especially under article 4 of their principles: "The data handled by the ICRC should be adequate and relevant to the purposes for which they are collected and processed. This requires, in particular, ensuring that the data collected are not excessive for the purposes for which they are collected and for compatible further Processing, and that the period for which the data are stored, before being anonymized or archived, is no longer than necessary."

20 The IOM Policy has a specific chapter on consent and notes the need to adopt a flexible approach given the nature of IOM's work and the contexts in which it operates, in particular it outlines different forms of consent, with emphasis on knowledge of the data subject being the bare minimum in situations such as emergency or conflict situations. See IOM Data Protection Manual, from p.21.

21 The ICRC Rules on Personal Data Protection make this point very explicit under article 1, section 3: "Wherever possible, Consent is the preferred basis for Processing Personal Data. However, because of the vulnerability of most of the beneficiaries of ICRC activities, and the nature of the organization's work in humanitarian emergencies, the ICRC may not be in a position to rely on this preferred basis for many of its Processing operations."

An integrated approach instructing risk assessment and risk mitigation along the data value-chain (see below), including both technical and organisational tools and methods, could generate an important model for the humanitarian sector. Including a benefit/risk assessment step in the instructions for data use would provide data handlers with a clear overview of how to check the balance between potential risk and intended benefits. Further, including the data value chain attached to risk assessment and mitigation, will provide data users (both inside and outside the organization) the required context to better understand the background and rationale of requirements.

C.5 VALUE -CHAIN

Describing the data value chain and risks/mitigation strategies in a policy step-by-step allows users of the policy to better understand why they are required to undertake the actions described in the policy, increasing the likelihood of implementation.

C.5.1. VALUE CHAIN AND RISK/BENEFIT DESCRIPTION

Most of the data policies reviewed are structured chronologically according to the typical data lifecycle leading roughly from collection through use of the data to dispatch of results. The layout of the HDX data policy is an example of a data policy with such a structure. For most of these policies, this value-chain is not made explicit, but rather implied in the structure of the policy governing it. USAID's Privacy Program lays out the data life cycle as follows: *"All USAID Managers must consider the information life cycle (i.e., collection, use, retention, processing, disclosure, and destruction) in evaluating how information handling practices at each stage may affect the privacy rights of individuals."*

Another noteworthy example is IOM's data policy, which contains a step-by-step assessment of risk and mitigation along the data life cycle. Noting that continually assessing the benefit/risk balance of data use throughout the life cycle of data processing is important because the balance may change. IOM lays out the 'data processing life cycle', noting that the data flow through its organisation is not a linear process, but rather that a variety of actions can be taken in parallel, such as research and transfer of the data.

C.5.2. DATA PROCESSING

More than half of the policies analysed (9 out of 17) contained a description of what was understood as 'data processing'. IOM defines data processing as "[t]he means by which personal data are collected, registered, stored, filed, retrieved, used, disseminated, communicated, transferred and destroyed." As for the description of data types, the reason for which some policies lack such a description might be increased flexibility. However, the International Charter for Space and Natural Disaster provides an example of how an

accessible description of the transition from data to information can still ground the policy in practice: *"The term "information" means data that have been corrected and processed by the parties using an analysis program, in preparation for use in crisis management by one or more associated bodies in aid of the beneficiaries; it forms the basis for the extraction of specific products for use on location."*

C.6. PRINCIPLES AND LEGAL FOUNDATION

Introducing the principles and legal documents that have guided the formulation of a data policy further affects the provisions contained in the policy itself, allowing for easier interpretation and closer alignment of data-driven practices with fundamental principles.[22]

C.6.1. GUIDING PRINCIPLES AND NORMS REFERENCED

13 out of 19 data policies analyzed contained reference to principles, statutes and/or regulations. Such sources are generally mentioned in an introductory section to the policy document. In a smaller number of documents, the legal basis was added in specific provisions.

IOM is an intergovernmental organization and notes that its 13 data protection principles have been informed by relevant international instruments. In Annex A to the IOM Data Protection Manual it lists privacy and data protection legal instruments applicable at the time of developing the IOM policy. The privacy protection policy of USAID shows a high amount of references to legal and ethical guidance. Under section 'privacy framework', reference is made to the 'fair information practice principles', and throughout the document relevant legislation is mentioned–with links–to explain the background of provisions in the policy. For example, in their article on Privacy Impact Assessments, USAID notes:

> *"This section addresses USAID's policy requirements for the creation and maintenance of Privacy Impact Assessments (PIAs) as required by Section 208 of the E-Government Act of 2002 and OMB implementing guidance. USAID must conduct a PIA when it uses information technology (systems) to collect, use, maintain, or disseminate PII."*

The MSF policy sets out 9 principles in its policy statement. Further, MSF acknowledges under point 1 on page 6, that its data sharing practices will comply with applicable national and international law. LIRNEasia references applicable law differently, recommending under 'Remedies' that Mobile Network Operators should not engage in surveillance of their customers, except when they are required to do so under applicable law.

22 For further insights on the application of the Humanitarian Principles to the use of ICT's, see: Nathaniel A. Raymond and Brittany L. Card, "Applying Humanitarian Principles to Current Uses of Information Communication Technologies: Gaps in Doctrine and Challenges to Practice", Signal Program on Human Security and Technology Harvard Humanitarian Initiative, July 2015, available from [*http://hhi.harvard.edu/sites/default/files/publications/signal_program_humanitarian_principles_white_paper.pdf*].

After compiling all principles found in the documents analysed and condensing the principles found, those referenced most often are the following:

- Confidentiality
- Respect the privacy and dignity of data subjects
- Accountability
- Protection of Data
- Neutrality/Impartiality/Independence of the humanitarian organization
- Transparency
- Dignity and Respect for the Data subject
- Openness of data
- Legitimacy of the Data use
- Integrity of the Data

C.6.2. GUIDING PRINCIPLES AND NORMS EXPLAINED

Of those documents that contained reference to principles and policies, several explained the contents of these reference materials, whereas others merely reference the title or a brief description of relevant principles. For example, the USAID data program makes reference to the Fair Information Practice Principles, giving a brief description of their contents.

C.7. TOOLS AND PRACTICES

C.7.1. SPECIFIED TOOLS AND PRACTICES FOR POLICY IMPLEMENTATION

Including tools and practices in a data policy is fairly rare and done by less than a third of the policies analysed. The 5 policies that do specify tools and/or practices for implementation of the policy include the GSMA guidelines for example, in which use of the SHA-3 algorithm is mandated for the hashing process to anonymise the CDR data used. The IOM data policy employs a range of templates and checklists to be used when handling data subject to the policy.

In article 5, the HDX data policy mentions several options for participating organisations to contribute data. These options, too, may be considered tools and practices for policy implementation as they guide the behaviour of actors subject to the policy in such a way that adherence is stimulated.

C.7.2. DECISIONS INFORMED BY TOOLS AND PRACTICES SPECIFIED

Only 3 policies describe the decisions informed by the tools and practices they specify. For the GSMA–mentioned above under 7.2–this decision relates to the standard of anonymisation acceptable for use of the CDR's in the Ebola response. For HDX, providing options to contribute data to the platform guides the decision of how to do so, but this is not made explicit. Listing tools and practices such as these can help to foster both data use as well as enhancing the chances of adherence to the policy. USAID, on the other hand, does include a description of the decision its tool aims to inform, stating that, *"[e]mployees must protect PII, as well as other SBU information, against unauthorized access or disclosure by ensuring that only those people who have a clearly demonstrated need to know or use the information have access."*

C.8. ACCOUNTABILITY AND DESIGN

C.8.1. POLICY DRAFTING PROCESS

Few privacy and data protection policies we analysed described the process by which they were formulated. The Principles for Digital Development are an example of a document for which the drafting process is clearly stated. An introductory statement explains that the principles represent *"a concerted effort by donors to capture the most important lessons learned by the development community in the implementation of technology-enabled programs."* The Privacy and Trust Principles for the White House Precision Medicine Initiative were formulated by an interagency working group and informed by a series of roundtables, literature review, analysis of privacy policies and frameworks, and comment from the public.

The IOM Data Policy is the most notable exception as it describes in detail how the policy was developed, with the following explanatory statement in its foreword:

> *"By way of background, IOM conducted a survey of selected registration projects in 26 field offices in 2007. The survey illustrated that there was indeed a need to standardize the handling of personal data throughout the Organization. IOM's policy on data protection is informed by relevant international standards, in particular the core data protection principles as recognized by many States, and through research on policies and procedures in other organizations."*

The ICRC introduction also provides a relevant exception, as it too describes the process by which the policy was formulated.

Adding the drafting process of the policy helps explain why certain angles and approaches are chosen in the policy, and adds legitimacy to the document. For any organization seeking to use data, it would be valuable to work out a detailed description of how it developed its data strategy in an inclusive and

transparent manner. Such a description could include an explanation of the constituencies that were included in the development of the policy, and perhaps an indication of some of the key outcomes of stakeholder consultations in the introduction to the policy.

C.8.2. PARTICIPATORY OR USER DESIGN PRINCIPLES

None of the data policies reviewed contained an elaborate description of a process to receive input and feedback from data subjects or those that are supposed to abide by the policy once it is in place. The Precision Medicine Initiative does state that a public consultation was held, but no statement is made on the outcome of this publication.

Inclusive policy development will lead to policies that are both of higher quality, have more impact and a higher chance of implementation, as well as higher legitimacy.[23] Although multiple current data policy development processes include both an expert consultation as well as user participation, making this process explicit is not yet common. Acknowledging this process and relating its outcomes to policy decisions will strengthen any policy development going forward.

C.8.3. MONITORING AND EVALUATION MECHANISMS IMPLEMENTED

A third of the policies analysed described mechanisms for monitoring adherence to the policy itself. For IOM, this is described as follows:

> *"An independent body should be appointed to oversee the implementation of these principles and to investigate any complaints, and designated data protection focal points should assist with monitoring and training."*

This type of indication of an oversight mechanism to improve implementation of the policy is relatively rare, and useful if implemented accordingly. The precise moments or circumstances under which the monitoring body should be called upon must also be contained in the policy, to enhance its effectiveness.

C.8.4. DISPUTE RESOLUTION MECHANISMS

Only 3 of the policies analysed contained mechanisms for dispute resolution. Oxfam states that, all participants in their programmes are free to withdraw their consent and to discontinue their involvement in the data activity at any point. IOM's data policy similarly includes a right on the part of data subjects to withdraw their consent at any time and gain access to their personal data and states that policy developments and practices should be transparent and allow for complaint mechanisms. The EDPR

23 Capelo, L., Chang, N., Verity, A. (2013) "Guidance for Collaborating with Volunteer & Technical Communities". CreativeCommons. DHNetwork. communitiesofinterest.net. accessed 9.4.2014. Milner M, Verity A. (2013). "Collaborative Innovation in Humanitarian Affairs: Organization and Governance in the Era of Digital Humanitarianism." http://blogveritythinkcom.

contains the 'right to be forgotten', which is an elaborate mechanism by which individuals can demand removal of results regarding outdated, incorrect or irrelevant information from search engine indexes. The ICRC rules contain a full chapter on data subjects rights, with a right for data subjects to object to the processing of their data, contained in article 11, and a procedure for dispute resolution:

- "Article 13: Assertion of data protection rights by individuals
- Data Subjects may make a formal assertion of their data protection rights with the ICRC Data Protection Office.
- When it cannot settle an individual complaint itself, the ICRC Data Protection Office refers the matter to the ICRC Data Protection Commission. If a complaint is found to be justified, appropriate measures should be taken."

Typically, disputes will arise between the organisation that uses the data, and individual data subjects. Including these individuals in the design of the policy on that data use, and providing them with easy access to voice concerns regarding the way their data is used, will often prevent many of these disputes from arising. In case a dispute does nonetheless occur, establishing a fair and quick procedure for dispute settlement can signal that an organization takes such concerns seriously. Lessons learned in dispute settlement should feed back into further iterations of the policy and other elements of any organization's overall data strategy.

C.8.5. EXPLANATION OF ROLES AND FUNCTIONS

A majority of the analysed policies describes specific roles and functions. Typically, the role of 'data controller' or a similar title is given to persons specified in the policy, with a description of the authority and responsibilities of such a position. The policy of UNHCR provides an excellent example:

- **Data Controller**: *The UNHCR staff member, usually the Representative in a UNHCR country office, who has the authority to oversee the management of, and to determine the purposes for the processing of personal data.*
- **Data Processor**: *Any UNHCR staff member or other natural person or organization, including an Implementing Partner or third party that carries out processing of personal data on behalf of the data controller.*
- **Data Protection Focal Point**: *In principle, the most senior UNHCR protection staff member in a UNHCR country office or operation, who assists the data controller in carrying out his or her responsibilities regarding this Policy."*

In the USAID Privacy Program, guidelines are different specifically for those in 'manager' or 'privileged' positions. Creating these types of categories by which people working under a given data policy can easily determine which rules to abide by, may be an effective way to instruct data users, without cluttering the data policy with endless descriptions of specific roles.

The Charter for Space and Disaster assigns roles for 2 institutional bodies: *"3.3 The administrative, operational and technical coordination needed to achieve this cooperation shall be provided by a Board on which each party is represented and an executive Secretariat for implementation of the Charter."* The ICRC rules assign various roles and responsibilities to the ICRC Data Protection Office, including dispute resolution and monitoring and oversights of policy implementation.[24] The IOM policy describes various functions of data protection focal points to perform oversight and monitoring implementation of policy.[25]

A clear description of the tasks and responsibilities to enact the policy and monitor and enforce adherence to it, is one of the central aims of a comprehensive data policy. Assigning all actors involved with a clear role will prevent duplications in work and promote the implementation of data responsibility mechanisms as designed in the policy.

D. Takeaways

The burgeoning availability of data has led to growing expectations regarding the potential of data to prevent or mitigate crisis situations. At the same time, increased sharing and use of data are not without risks – including threats to individual privacy and security. A cross-sectoral body of 17 documents has allowed us to consider current policies and practices from around the world that promote responsible use.

The above includes the detailed results of our analysis along eight broad themes. In this concluding section, we include some additional lessons and takeaways. These are designed to help humanitarian organizations design their own policies and guidelines regarding the effective and responsible use of data.

1. Fragmentation and the need for leadership: In the course of this study, we identified fragmentation at multiple levels. There appear to be a lack of:

- **set of responsible data principles** that, when effectively used, can create a culture of data responsibility empowering humanitarian actors with the information they need in order to make decisions that save or improve people's lives. Instead, as our comparative assessment identified, various norms and values are used inconsistently, depending on the actor, system and/or purpose;
- **responsible data framework** comprising policies, practices, and tools that implement the accepted data principles and institutionalize responsibilities and procedures.

2. Coordination and leadership

- Despite a global consensus on the need for data governance, actors and stakeholders are still debating the precise contours of such governance. We identified several organisations whose policies are more advanced in their thinking and scope on this issue, both in terms of the detail with which

24 For a description of the ICRC Data Protection Office, see: [https://www.icrc.org/en/document/icrc-data-protection-office].
25 IOM Data Protection Policy, p. 98.

potential risks are described, as well as the elaboration of new ways to mitigate these risks. These include UNHCR, IOM, ICRC and USAID. In addition, the work done at UN Global Pulse and its Privacy Advisory Group aims to provide new, workable approaches that could be applied widely across the humanitarian community. The work done by these organizations offers a useful template for other organizations currently considering using data or increasing their use of data for humanitarian work.

3. The need to consider trade-offs: Designing and drafting policy is often a process of tradeoffs. Some of the choices that organizations are likely to face in drafting a data use policy include:

- detail versus vagueness, which can determine the level of compliance or increased reliance on procedure;
- specificity versus flexibility, which determines the extent to which a policy regime is enabling or constraining;
- simplicity versus complexity, which affects the extent to which policies and their underlying rationales are properly understood (and thus achieve buy-in).

In many cases, a hybrid approach can be taken that attempts to combine the best elements of these various choices. Such an approach should ideally emerge as the result of a better understanding of the needs of users and the intended audience.

4. The need to determine responsibilities and roles: any effective responsible data use policy must include a clear sense of individuals or groups that are accountable or tasked to oversee the policy's implementation and enforcement. For this reason, any data policy should clearly determine and indicate responsibilities and roles. Some policies we have examined centralize this function, while others take a more decentralized approach. Our general sense is that a decentralized or distributed approach is more effective, allowing a broader community to take part in decision-making, but the ultimate structure and delineation of duty must be based on each organization's own needs and feedback from its user community and other stakeholders.

5. The need for innovation: solving 21st century challenges using 20th century tools and procedures will never be effective. Virtually all the policies identified here are still using what we may think of as "traditional" means to govern. Responsible data use approaches should instead consider innovative methods and processes that have the potential to increase both legitimacy and effectiveness. In particular, in drafting a data governance frameworks, organizations may want to consider:

- co-creation methods to collaboratively draft policies and approaches;
- behavioral economics or nudging approaches to inform particular choices and behaviors;
- decision trees and other expert systems to guide implementation of policies and principles;
- peer review mechanisms to encourage a distributed management system.

6. The need to use clear language: Good policies are easy to read. Among our examples, IOM's policy exemplifies the use of plain language. In this context, it is interesting to note that some of the reviewed policies made substantial use of visual elements, which adds to the usability and comprehensibility of the policies.

7. The need to gather evidence on what works. In order to be truly effective (and responsible), data policies should themselves be data-driven. To date, little is known about what actually works, and under what conditions. This report represents an initial attempt to address this shortcoming. More generally, data policies should be developed in an iterative and evidence based manner—testing and adjusting approaches based on impact and user feedback. Addressing the mechanisms and approaches to institutionalize the learning of what works or does not work, and how these lessons feed into following iterations of the policy itself, is recommended.

8. In conclusion, the lack of a responsible data framework for the humanitarian space presents a substantial risk: Without an adequate data governance framework and established policies that are supported and implemented throughout its organization, actors in the humanitarian space put not only their own organization at risk, but more importantly, their beneficiaries as well. Specifically, without a set of core principles, translated to standards, practices and operational guidance to mitigate the risks of collecting, processing, and using data, the use of digital data in a humanitarian setting can result in a variety of risk scenarios, potentially causing greater harm to already crisis-affected individuals and their local communities, governments, and societies.

APPENDIX 1. REPOSITORY OF DOCUMENTS REVIEWED
APPENDIX 2. ANALYSIS TABLE

Appendix 1. Repository of Documents Reviewed

To benchmark and compare data policies within international organisations, we have curated the following policy documents for further analysis.

Medicins Sans Frontieres–*Data Sharing Policy, 2013*

Organization type: NGO

Sector and purpose for the use of data: Humanitarian Assistance in countries with endemic diseases. Data is used for researchers working in public health.

Document type: Report

Document background: Document was written to provide a data sharing policy for MSF related operations.

Link to document: *http://www.msf.org/sites/msf.org/files/msf_data_sharing_policy_final_061213.pdf*

Oxfam–*Responsible Program Data Policy, 2015*

Organization type: Oxfam, NGO

Sector and purpose for the use of data: Famine and poverty relief. Data is used for Oxfam reports.

Document type: Report

Document background: Document was designed to be a forward-looking attempt to prepare Oxfam for the "data revolution".

Link to document: *http://policy-practice.oxfam.org.uk/publications/oxfam-responsible-program-data-policy-575950*

UN Population Fund–*Information Disclosure Policy, n.d.*

Organization type: UN Agency

Sector and purpose for the use of data: Poverty reduction and minority rights, Data is collected to support prorammes.

Document type: Website

Document background: Policy was designed for their data collection from data users in the field.

Link to document: *http://www.unfpa.org/information-disclosure-policy#*

UN-OCHA–*Humanitarian Data Exchange–Terms of Service, n.d.*

Organization type: UN Agency

Sector and purpose for the use of data: Humanitarian and Crisis Response. Hub for sharing data across the humanitarian community.

Document type: Website

Document background: Hub for data sharing. Document lays out terms and conditions for the data base.

Link to document: *https://data.hdx.rwlabs.org/about/terms*

UNHCR–*Policy on the protection of Personal Data of Persons of Concern to UNHCR, 2015*

Organization type: UN Agency

Sector and purpose for the use of data: Humanitarian response. Data is used for UNHCR programmes.

Document type: Report

Document background: Report was written to lay down the rules and principles to data processing and collection.

Link to document: *http://www.refworld.org/docid/55643c1d4.html*

LIRNEasia–*Draft Guidelines for Third-Party Use of Big Data Generated by Mobile Network Operators*

Organisation type: Research Institute

Sector and purpose for the use of data: Various, draft insights are meant to be broadly applicable

Document type: Draft principles

Document background: Principles were drafted to focus and enhance discussion around the principles for the use of mobile telephony provider network data

Link to document: *http://lirneasia.net/wp-content/uploads/2014/08/Draft-guidelines-2.2.pdf*

GSMA–*Guidelines on the protection of privacy in the use of mobile phone data for responding to the Ebola outbreak*

Organisation type: International association (of mobile telephony providers)

Sector and purpose of the use of data: Strengthen the effort against the 2014 Ebola outbreak

Document type: Guidelines

Document background: Drafted to facilitate the use of cell phone data in the Ebola response

Link to the document: *http://www.gsma.com/mobilefordevelopment/gsma-guidelines-on-the-protection-of-privacy-in-the-use-of-mobile-phone-data-for-responding-to-the-ebola-outbreak*

White House Precision Medicine Initiative–*Privacy and Trust Principles*

Organisation type: US Government

Sector and purpose of the use of data: Enhance health research by tapping into new data sources

Document type: Principles

Document background: Drafted to guide the Precision Medicine Initiative in its exploration of the potential of new data sources for health research

Link to the document: *https://www.whitehouse.gov/sites/default/files/microsites/finalpmiprivacyandtrustprinciples.pdf*

UN Global Pulse–*Privacy and Data Protection Principles*

Organisation type: UN unit

Sector and purpose of the use of the data: International development and humanitarian response

Document type: Principles

Document background: Drafted to guide Global Pulse's work, together with the UN Global Pulse Privacy Advisory Group (http://www.unglobalpulse.org/data-privacy-advisory-group)

Link to the document: *http://www.unglobalpulse.org/privacy-and-data-protection-principles*

UN Office for Outer Space Affairs–*International Charter for Space & Major Disasters*

Organisation type: UN agency

Sector and purpose of the use of the data: Freeing up the flow of data in disaster situations

Document type: International Charter

Document background: Drafted to facilitate sharing satellite imagery during emergencies

Link to the document: *https://www.disasterscharter.org/web/guest/text-of-the-charter*

UN-OCHA–*Humanitarian Principles*

Organisation type: UN agency

Sector and purpose of the use of the data: International Organisation, humanitarian response

Document type: Principles–humanity, neutrality, impartiality and independence

Document background: Principles guiding humanitarian actors

Link to the document: *https://docs.unocha.org/sites/dms/Documents/OOM-humanitarianprinciples_eng_June12.pdf*

Principles for Digital Development–*Living Guidelines for Technology-Enabled Programs*

Organisation type: Collective of international agencies and donors

Sector and purpose of the use of the data: -

Document type: Principles

Document background: Principles guiding the use of digital tools for development

Link to the document: *http://digitalprinciples.org/*

EU Regulation: *On the protection of individuals with regard to the processing of personal data on the free movement of such data (General Data Protection Regulation).*

Organization type: Intergovernmental organization

Sector and purpose of the use of the data: Not specified

Document type: Legislation

Document background: EU Legislation

Link to the document: *http://ec.europa.eu/justice/data-protection/document/review2012/com_2012_11_en.pdf*

International Organization for Migration–*Data Protection Policy*

Organization type: International Intergovernmental Organization

Sector and purpose of the use of the data: protection of personal data of IOM beneficiaries in the migration context

Document type: Data protection policy

Document background:- Manual comprised of Principles and Guidelines

Link to the document: *http://publications.iom.int/system/files/pdf/iomdataprotection_web.pdf*

UNICEF–*Information Disclosure Policy*

Organization type: UN agency

Sector and purpose of the use of the data: Crisis / Children

Document type: Data policy

Document background:-

Link to the document: *http://www.unicef.org/about/legal_58506.html*

USAID ADS Chapter 508–*Privacy Program*

Organization type: US Government

Sector and purpose of the use of the data: Crisis response

Link to the document: *https://www.usaid.gov/sites/default/files/documents/1868/508.pdf*

International Committee of the Red Cross–*Rules on Personal Data Protection*

Organization type: NGO

Sector and purpose of the use of the data: Crisis response

Document type: Personal Data Policy

Document background: New rules on the use of personal data, to keep in pace with recent technological developments.

Link to the document: *https://www.icrc.org/en/publication/4261*

Appendix 2. Comparative Table

	MSF	UNHCR	DP	HP	SPACE	UNGP	PM	DGM	HDX	UNFPA	OXFAM	GSMA	EUREG	USAID	UNICEF	ICRC
SCOPE																
1.1. Clarification of goals and objectives	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
1.2. Description of context	✓	✓						✓	✓	✓		✓	✓	✓	✓	✓
VALUE PROP																
2.1. Value proposition description					✓		✓		✓	✓	✓	✓	✓		✓	✓
2.2. Data impact measurement			✓													
DATA																
3.1. Data types	✓				✓		✓	✓	✓	✓	✓	✓		✓	✓	✓
3.2. Purpose and anticipated value audit	✓	✓							✓	✓	✓			✓	✓	✓
RISK MANAGEMENT																
4.1. Risk assessment processes		✓	✓				✓		✓		✓		✓	✓		✓
4.2. Risk mitigation and response		✓				✓	✓	✓	✓		✓	✓	✓	✓		✓
VALUE CHAIN																
5.1. Value chain and risk/benefit description									✓		✓	✓		✓		
5.2. Data processing		✓			✓		✓		✓		✓	✓	✓		✓	
PUBLIC RELATIONS																
6.1. Guiding principles and norms referenced	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
6.2. Guiding principles and norms explained			✓			✓	✓		✓	✓	✓		✓	✓		
TOOLS AND PRACTICE																
7.1. Specified tools and practices for policy implementation									✓	✓	✓	✓		✓	✓	✓
7.2. Decisions informed by tools and practices specified												✓		✓		✓
ACCOUNTING AND DESIGN																
8.1. Policy drafting process			✓		✓		✓	✓	✓	✓	✓		✓	✓		
8.2. Participatory or user design principles		✓	✓		✓		✓							✓		
8.3. Monitoring and evaluation mechanisms implemented		✓					✓		✓		✓		✓	✓	✓	✓
8.4. Dispute resolution mechanisms									✓		✓		✓			
8.5. Explanation of roles and functions	✓	✓			✓			✓	✓	✓	✓		✓	✓		✓