STUDY

Beyond the Numbers:

Balancing Innovation, Ethics, and Impact

UPDATED PANORAMA

of the program data practices

and needs of CSOs

CartONG

SEPTEMBER 2024

1. INTRODUCTION

CARTONG

Created in 2006, CartONG is a French H2H/support NGO specialized in Information Management. Our goal is to put data at the service of humanitarian, development and social action projects. We are dedicated to improving the quality and accountability of field activities, in particular through better needs assessments and monitoring and evaluation. We act as a multidisciplinary resources and expertise center, accompanying our partners' strategies and operations. Our staff and volunteers also support the community as a whole by producing documentation, building capacities and raising awareness on the technical, strategic and ethical challenges of digital technologies.

AUTHORS

This study was carried out in the framework of the "Strengthening Information Management within francophone CSOs" initiative led by CartONG and co-financed by the French Development Agency (AFD) since 2019- the project that makes available the IM resource portal and its learning corner. For more information see the article published on our website: https://www.cartong.org/en/2024/05/31/pstrengthening-francophone-csos-in-data-management-phase-2-2/, write to: renforcement-osc@cartong.org or through our feedback form.

This study was written by the CartONG team - in particular Martin Noblecourt and Maeve de France, with the support of Léonie Miège, William Natta, Pauline Michavila, Berhudan Mustafa and Christophe Rodier.

The study was translated from English to French by Marie Parisot with support from the CartONG team. The illustrations were produced by Sophie Aumas and William Natta.

ACKNOWLEDGMENTS

CartONG thanks the partners of the project and the members of the "francophone Information Management Community of Practice" for the richness of the exchanges around these themes over the last nine years.

A big thank you to all the respondents of the survey and to the professionals of the humanitarian and development sectors who accepted to be interviewed for this study or participated in the various workshops as well as Groupe URD, Bioforce and the CyperPeace institute for their review of certain sections.

This publication is supported by the French Development Agency (AFD). Nevertheless, the ideas and opinions presented in this document do not necessarily represent those of the AFD.

This study is made available under the terms of the Creative Commons Attribution 4.0 International License.

: https://creativecommons.org/licenses/by/4.0/deed.fr

Readers are encouraged to use the contents of this study for their own publications, as long as they duly refer to it when it is mentioned (quotation, excerpt, name of publication, etc.). For online use, we request that the link of the publication on the CartONG website be used.

TABLE OF CONTENTS

1. INTRODUCTION 2

2. EXECUTIVE SUMMARY 6

3. RATIONALE 8

4. METHODOLOGY 11

4.1. SCOPE AND TERMINOLOGY 11

4.2. SOURCES 12

4.3. LIMITATIONS 13

5. USE OF DATA: WHAT CHANGES AND PERSPECTIVES OF CHANGE? 15

5.1. DATA FOR WHAT AND FOR WHOM? 15

5.2. DATA QUALITY: AN EVERLASTING QUEST? 21

6. A FIELD UNDER PROFESSIONALIZATION 28

6.1. “PROGRAM DATA”, A BETTER UNDERSTOOD AND RECOGNIZED CONCEPT 28

6.2. GROWING SKILLS 30

6.3. THE ORGANIZATIONAL LEVEL’S CHALLENGE 34

6.4. A PROFESSIONALIZATION CONNECTED TO DATA LITERACY 37

7. NEEDS: SOLUTIONS & RESOURCES COMPATIBLE WITH SECTOR SPECIFICITIES 40

7.1. ENSURING TECHNICAL SOLUTIONS… 40

7.2. …AND RESOURCES… 42

7.3. …THEN CHANNELLED THROUGH VARIED LEARNING FORMATS ... 47

8. LE RISQUE D’UN SYSTEME A 2 VITESSES POUR LES DONNEES HUMANITAIRES 49

9. A CHANGING FIELD: CURRENT AND FUTURE STAKES 57

CHAPTER METHODOLOGY AND INTRODUCTION 57

9.1. DATA GOVERNANCE & LOCALIZATION 59

9.2. ACCOUNTABILITY TO AFFECTED POPULATIONS (AAP) 63

9.3. DATA-DRIVEN DECISION MAKING 69

9.4. CONSISTENT STRATEGIES 73

9.5. DATA LITERACY 77

9.6. RESPONSIBLE DATA MANAGEMENT 80

9.7. CYBERSECURITY 85

9.8. STANDARDIZATION, DATA SHARING & OPEN DATA 91

9.9. LEGAL AND CONTRACTUAL ENVIRONMENT 93

9.10. RESPONSIBLE USE OF AI 98

9.11. SUSTAINABLE TECHNOLOGIES 103

9.12. MASTERED DIGITAL STAKES 106

10. RECOMMENDATIONS 113

10.1. ALL ACTORS: PROMOTE COLLECTIVE STRUCTURING, LEARNING AND CRITICAL THINKING 113

10.2. CSOS: CONTINUE PROACTIVELY INVESTING THE PROGRAM DATA BRANCH 115

10.3. NETWORK HEADS: SUPPORTING THEIR MEMBERS IN THE IMPROVEMENT OF THEIR PRACTICES 116

10.4. DONORS AND UN AGENCIES: SUPPORT AND ENCOURAGE THE PROGRAM DATA UPTAKE OF CSOS 117

10.5. FOR SPECIALISED SUPPORT CSOS: BETTER ASSIST CSOS NAVIGATE THE DATA STAKES OF THE DAY 120

11. CONCLUSION 122

2. EXECUTIVE SUMMARY

Data management plays a central role in the evolutions of international aid, at the crossing of efficiency, accountability and innovation. Following our initial work "Program Data: The silver bullet of the humanitarian and development sectors" in 2020, CartONG continues explorng in this study recent evolutions as well as less visible consequences of the "datafication" of humanitarian CSOs' work. This new study endeavours a larger review: both in terms of geography (with an effort to open globally to more cultural contexts, in particular from the Global South) and in terms of type of actors (with more peeks in the world of United Nations, governments, and the private sector).

Our work builds on the triangulation of a survey (filled by 62 participants from 54 organisations), 13 semi-structured interviews, 4 focus-groups interviews and a documentary review of more than 100 sources. Despite some limitations (number and diversity of respondents still not statistically representative, in particular from the "Global South"), our study brings a rare synthesis of the state of the art of data management among humanitarian CSOs.

The study begins by reviewing changes and perspectives of changes in data use, starting with the root question of why we produce and use data, and for what and whom. We comment on the relative lack of evidence of impact of data on policy-making, and a data collection still mostly driven by upwards accountability. At the same time, critical flaws in data quality, such as biases in data and underrepresentation (e.g. language, gender), lack of quality control during data collection, and persistent under-use of qualitative and secondary data.

We then examine the state of professionalisation of the sector, with an uptake of the appropriate of "program data" management concept, growing skills on all steps of the data management cycle – with some persisting differences between international and local CSOs – and significant needs to improve methodological and organizations aspects rather than purely technical skills, and a progress of data literacy among non-specialists that still remains insufficient. We also review the current needs of CSOs in terms of solutions and resources. With the profusion of technical solutions, the challenge seems more to be identifying the proper solution aligned to the CSO's needs (including assessing needs); getting sufficient support and funding to properly implement it; and having access to learning resources and opportunities.

All these evolutions however need to take into account risk of a 2-tier humanitarian data system, with a growing gap between a few global actors (UN, major CSOs,

donors, Global North NSOs…) and most of the others (most CSOs including the smaller international ones and local/national, Global South NSOs…). Without corrective action, smaller and local actors might enter in a vicious circle of suffering technical evolutions and increasing accountability requests without resources to appropriate / control them.

Our study then examines the current and future stakes of humanitarian data management highlighted by our survey respondents: first, data governance & localization, and accountability to affected populations, two topics that imply a deep transformation of our vision of data management. Then data-driven decision making, consistent strategies, and data literacy, the three components of a successful onboarding of data in CSOs. We then cover the connected topics of responsible data management and cybersecurity, two critical risks for the sector. We finally discuss standardisation, data sharing & open data, the legal and contractual environment, responsible use of AI, sustainable technologies, and mastered digital stakes, all stakes that imply adapting to evolutions while preserving the capacity of action of CSOs.

The study concludes with recommendations for the humanitarian system, for CSOs, for network heads, for donor & UN agencies, and for specialized support CSOs, in terms of promoting collective learning and critical thinking, continue proactively invest in program data and improving practices, and how to assists CSOs the data stakes evoked in the report.

3. RATIONALE

With the multiplication of crisis in a world facing threats of diverse nature (resurgence of armed conflicts, impact of climate change, global recession…) the need for humanitarian aid is constantly growing, estimated from 38,5 billion USD in 2020 to 56 billion USD in 2023, more than tripling in 10 years[1]. At the same time, the resources available are still scarce, with a global coverage of coordinated plans around 50% over the past years. In this context, the humanitarian sector must constantly improve its efficiency, accountability and innovate to adapt to these challenges. Data management plays a pivotal role in these evolutions: allowing accurate assessment of needs, (almost) live monitoring and rational impact evaluation are now mandatory for humanitarian organizations, whatever their size, location and level of structuring. Programme data, the "silver bullet" of the sector, has become a key decision-maker for an industry that seeks to be "data-driven".

While data management provides immense services to our sector, its growing strategic role also raises critical question on its impact. While the necessity to deliver highly needed services to affected populations around the world leaves little room for a collective capacity to question our practices, more and more actors develop a reflective perspective on them. Including the consequences of "datafication" of our day-to-day work, the associated risks of always more digital processes, and ultimately the risk of a bureaucratization, dehumanization or diversion from humanitarian principles.

CartONG's mission statement is precisely to support humanitarian actors on tackling the technical, strategic and ethical challenges of digital technologies, in particular data technology: this is why we will try to measure these opportunities and risks through the study you are reading, building on the collective intelligence of our peers.

A key milestone for us was reached in 2020 with the publication of "Program Data: The silver bullet of the humanitarian and development sectors?"[2], one of the first studies to examine, as a whole, the practices and needs of (francophone) CSOs in terms of data management. The study was funded by the French Development

[1] "Humanitarian aid contributions 2024", OCHA Financial Tracking Services, 2024
[2] "Program Data: The silver bullet of the humanitarian and development sectors? Panorama of the practices and needs of francophone CSOs", CartONG, September 2020

Agency, as part of our "Strengthening NGO practices in Data Management" initiative (phase 1, 2020-2023[3]), hence this geographic focus. As we start a second phase of this project[4], it seemed necessary to update this study both as food for thought for the sector, and as guidance for the activities of our initiative.

This study therefore explores two main axes: first, analyse the impact of recent evolutions of technology for the sector (Artificial Intelligence, widespread diffusion of some tools, etc.); second, to continue investigating less visible consequences of the "datafication" of humanitarian action (governance in a changing landscape with localization, evolving humanitarian organisations, anticipatory action, privacy and rise of cyber threats, links with the private sector, etc.). These questions will be explored via thematic focuses through the study.

Compared to our 2020 study, this update had the advantage of building on a baseline. However, the changes in the humanitarian sector as well as the feedback received from various stakeholders (both as part of the evaluation[5] of our previous phase and through the discussions leading to this study), conducted us to expand slightly the angle and subject of this second study.

Firstly, while our project focuses on support to francophone actors, we have fully faced the fact that it would not make sense to study their positioning without considering their wider environment. Hence a larger review: both in terms of geography (with an effort to open globally to more cultural contexts) and in terms of type of actors (with more peeks in the world of United Nations, governments, and the private sector).

Secondly, we have tried to integrate fully the perspective of Global South actors in this study. Another more recent study of ours, "Changing the outlook: for a local approach to data"[6], published in January 2024 has allowed us to largely explore this question. We have endeavoured to embed here the learnings from this study to try as much as possible to present a global vision on the question, not limited to Global North international CSOs.

[3] "Strengthening NGO practices in Data Management (phase 1)", CartONG, 2023
[4] "Strengthening francophone CSOs in data management: phase 2", CartONG, May 2024
[5] [In French] "Évaluation finale du projet « Renforcer la gestion des données programmes des OSC", Key Aid Consulting, October 2022
[6] "Changing the outlook: for a local approach to data", CartONG, January 2024

While many stakes remain unchanged after 4 years, this study therefore aims to measure (when possible) the evolutions to our foundational 2020 work, but also to expand the scope to angles that connect to the wider changes of the humanitarian system. As for other fields such as Monitoring & Evaluation or Information Technology, (program) data management is now well identified as a key field of work. However, knowing this is not enough: our study also aims at provoking debates and stimulate changes among different level of stakeholders of data, to tackle the various challenges identified.

4. METHODOLOGY

4.1. SCOPE AND TERMINOLOGY

Similarly to our original study, this work focuses on the Civil Society Organizations (CSOs) of the humanitarian aid and international development sector – shortened as "humanitarian sector" in the study (we have not re-used the "Humanitarian Aid and International Development" or HAID acronym used in 2020 as it is not widely used in the sector, but it is to be understood as a synonym). As explained in the Rationale, we however expanded the scope to try to cover not only francophone but also more globally the stakes impacting most of the CSOs intervening in the "aid industry".

While the definition of what a CSO is can be debated, we include in it non-profit and non-governmental organizations, whatever their size and country of origin – in other words both "International" and "Local" CSOs. We will similarly not open the debate on the definition of "civil society", whose shape varies depending on contexts and cultures.

Our definition of CSOs thus excludes United Nations agencies, International Organizations, governmental actors, and for-profit companies. While these will all show up in our research through their numerous interactions with CSOs – as prescribers, partners, sometimes even adversaries – our study does not pretend to portray a comprehensive picture of the humanitarian sector. In fact, some of the assessments on the level of maturity discussed here do not apply to major organizations such as the UN or some "top tier" CSOs, as discussed in chapter 8. Having more capacity in terms of data in general, these organizations also tend to produce more learning on their practices[7]. It thus seemed relevant to focus on actors with less resource for self-reflective work, namely CSOs.

The definition of what is "local" is also a debate in itself[8]. We will use in this study "local" CSOs to define all organizations that have mostly a local or national presence (thus including "national" CSOs), and that don't have the scope and leverage of international CSOs. Since the localization question is deeply

[7] Cf. for example the work of the United Nations Department of Economic and Social Affairs' Statistics; the World Bank's DataBank and other projects; or the Red Cross movement's numerous data initiatives.

[8] "A more localised aid system: current status and discourse", Vijayalakshmi Viswanathan, ALNAP Essential Briefings for Humanitarian Decision-makers (EXplain), June 2023

connected to decolonization, it is to be understood that “local” CSOs are usually less in the power seat of the governance of the sector, as illustrated in “Changing the outlook: for a local approach to data”[9]. We acknowledge the fact that using the “local” term can be reductive of the role and impact of national organizations, but used it out of simplicity and to relate to the global localization agenda.

4.2. SOURCES

This study is based on 3 main sources of information that were triangulated during the analysis: an online quantitative and qualitative survey of CSOs, semi-structured interviews & collective focus-group type workshops, and a secondary literature review.

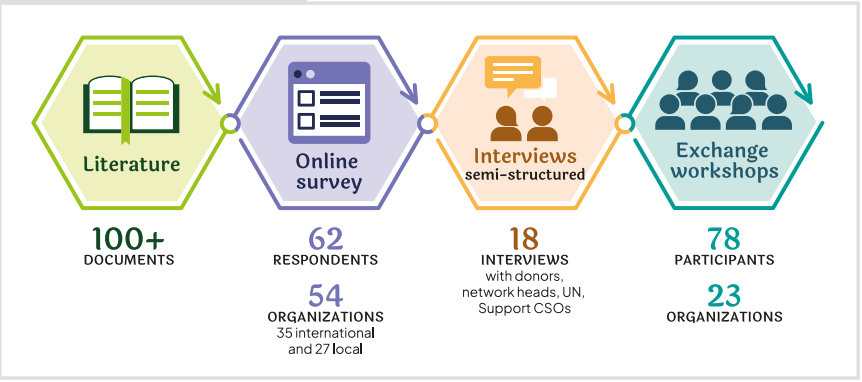
The survey (form in Annex 6) was disseminated between June and July 2024 by CartONG, with various partners relaying it also. After filtering out-of-scope respondents, it amounted to 54 organisations and 62 responses (vs.50 in 2020 – cf. list in Annex 4), 66% of which were data specialist profiles filling a long version of the survey (the other 34%, who were “generic” humanitarian professionals, filled a shorter form). They work from a variety of thematic sectors inside the humanitarian and development field (all the 15 sectors identified were represented with between 2 to 11% of respondents in each) – in fact, 53% of them working both on humanitarian response and development activities. Our panel also had an interesting diversity in size, with 24% of very small CSOs (yearly turnover < 250k €), 9% of small to medium CSOs (turnover 250k to 2M €), 15% large CSOs (turnover 2M to 10M €) and 39% very large CSOs (turnover > 10M €).

The survey was complemented with 13 semi-structured interviews with key stakeholders (list of interviews available in Annex 2 and semi-structured grid in Annex 3). We focused in particular on actors interacting with CSOs: donors and UN agencies, network heads, and support actors (from the H2H Network[10] in particular). To these interviews were added 4 focus-group type workshops: during the inter-CSO francophone exchange day; with the members of the francophone IM community of practice; with data specialists from H2H Network members; and a series of internal ones with the CartONG team (in June/July 2024, more information in Annex 5).

[9] “Changing the outlook: for a local approach to data”, op. cit.
[10] “About us”, H2H Network

Finally, we completed with a review of more than 100 documents, either pre-identified by our team during the 2021-2024 period as relevant for future research, or through custom research on specific topics we wanted to investigate further (bibliography available in Annex 1). Consistently with our extended focus, we did not emphasise any specific geography for this literature, but it ended up being anglophone to a great extent, reflecting the dominance of this language in both the humanitarian and technology sectors.

Data sources for the study



4.3. LIMITATIONS

Conducting a landscape analysis on program data management uses by humanitarian CSOs is so wide an exercise that being comprehensive would be an illusion. Our study is not exempt from this flaw: while we reached a larger representativity of our sources than the 2020 version, it is still limited in scope. Both the respondents to our survey and our interviewees were mostly coming from CartONG's extended network (for instance 74% of the respondents of the survey filled it in French and among the CSOs represented 44 out of 62 were based in Francophone countries, inc. 21 from France itself). The overall number of organizations either surveyed or interviewed, 58, while significant, cannot also claim to be statistically representative of a whole sector.

Our study still has a limited representation of "Global South" (or Global Majority) actors: the survey represented a great progress with 44% "local" actors filling it, but interviews were much more difficult to organize remotely, and the literature on the topic is almost exclusively originating from Global North actors. We nevertheless tried to give a global picture of the sector, presenting the position (and power struggles) of all actors, regardless of their location. As for most

localization efforts in the sector, following our dedicated publication already mentioned[11], the voice of Global South actors is gaining better representation in our work, but there is still a long path to go. We can in particular be satisfied to have respondents from 13 Global South countries to our surveys, balancing (even if often only one organization filled it per country) in these fields of interventions of the "humanitarian" sector the voice of international agencies, and providing some much needed diversity of perspectives.

Regarding the motivations of our research participants, they haven't received any technical or funding compensation for their participation (except project associates who proofread it), thus avoiding any bias. CartONG's own positioning of "think and do tank" is of course not neutral itself: as a Global North-based organization with mostly Western staff, as an Humanitarian2Humanitarian service provider, and as the promoter of a certain vision of data management for the sector.

Updating process: while it was not certain in 2020 CartONG would have the resources and capacity to update our original study, time has proved the interest for the topic (and hopefully our contribution!) has made it possible. We again cannot commit for a future update but encourage you to send comments, feedback or suggestions for next research to the following email address: renforcement-osc@cartong.org. We will take them into account in our future research work, and in particular if we re-update this study!

[11] "Changing the outlook: for a local approach to data", op. cit.

5. USE OF DATA: WHAT CHANGES AND PERSPECTIVES OF CHANGE?

5.1. DATA FOR WHAT AND FOR WHOM?

We chose to start this updated study by going back to the root question of the topic: why do we produce and use data? In other words, for what and for whom? In the 2010s, the "Data Revolution" narrative was built on an assumption that data would flow freely and be used for widespread data-driven decision-making. While there are countless examples of case studies demonstrating the impact of data on the efficiency of humanitarian work in many specific situations, we wanted to question here this global assumption.

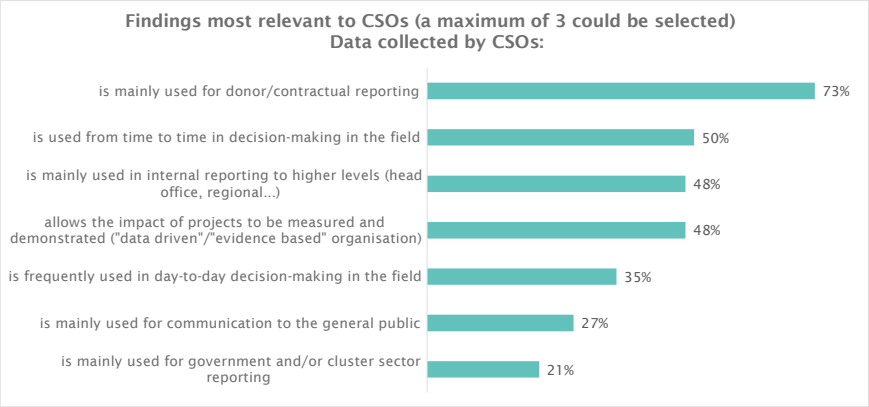
On the one hand, it is clear that the volume of data generated worldwide has boomed (from 4.4 zettabytes in 2013 to 175 zettabytes by 2025), with a diversification of the sources of this data (satellite, smartphones, citizen-generated projects…). But on the other hand, there is still no strong evidence that data is used to inform and drive policies at a global level[12]. Although it was a core assumption of the Sustainable Development Goals, several case studies (including a 2022 survey of National Statistics Officers) are also not conclusive in that direction.

> THERE HAS BEEN LIMITED CONSOLIDATED EVIDENCE OR GLOBAL ANALYSIS OF HOW POLICYMAKING PROCESSES IN THE CONTEXT OF SDGS HAVE BENEFITED OR CHANGED DUE TO IMPROVED STATISTICAL DATA PRODUCTION. – TRENDS

The little research available shows that this link between data and decision-making is, as one could expect, much more complex (as we had touched on also in our 2020 study), requiring understanding individual, organizational and system-level dynamics and working on buy-in by decision makers, as well as building capacities to analyse and use the data.

[12] "Testing the Assumptions of the Data Revolution", Thematic Research Network on Data and Statistics, 2024

So why do we collect data? If we focus on CSOs, the respondents to our survey in a large majority (74% – and 80% of international CSOs!) identify the need for upwards accountability (to donors) as their main driver to collect data. Using this data for operational monitoring and decision-making, as well as impact evaluations are well behind, with only half of the organizations mentioning it – at the same level as internal reporting of the organization. While this result must be nuanced as two-thirds of our respondents are data specialists – who might be less optimistic on the concrete impact of data than their programme colleagues – there are systemic causes and consequences of this situation that we will examinate later in this study (chapters 9.3 and 9.4).



Les participant·es à notre journée d'échange francophone de juin 2024 ont témoigné qu'il arrive encore souvent dans les opérations humanitaires qu'une décision soit prise, et seulement ensuite la justification trouvée par les données[13]. Peu d'acteur·rices ont le temps ou les ressources nécessaires pour analyser correctement les données : l'organisation spécialisée dans la qualité des données et la redevabilité Ground Truth Solutions nous a donné l'exemple d'une enquête de perception sur les transferts monétaires, dans une zone spécifique au Nigéria, où les personnes enquêtées ont exprimé une préférence pour les coupons. En raison de l'acceptation très large du transfert monétaire comme approche de référence, les acteur·rices de la réponse ont utilisé cette information comme point de départ pour une discussion, ce qui s'est avéré une décision intelligente car les

[13] Séminaire sur la Gestion des Données programmes, juin 2024

recherches qualitatives ultérieures ont montré un tableau plus nuancé – les gens n'aimaient pas les *spécificités* des transferts monétaires dans la région par les ONG, pas le principe. La solution était plutôt d'améliorer les transferts monétaires plutôt que de revenir aux coupons… mais les acteur·rices auraient-ils pris le même soin et le même temps pour vérifier cette information s'ils n'avaient pas « attaqué » une approche aussi reconnue que les transferts monétaires ? Cette question connecte à des mécanismes psychologiques tels que les biais de confirmation, qui sont eux-mêmes étudiés par les spécialistes[14].

Une autre façon d'évaluer cette question est de demander aux spécialistes de la donnée eux-mêmes : lors d'un atelier organisé avec des organisations spécialisées en gestion de données du réseau H2H, sur un panel de 18 expert·es de 10 organisations, elles et ils ont estimé à 2,9 sur 5 leur capacité à évaluer l'impact des données/outils/supports fournis (avec une part importante mettant des scores très faibles) :

> JE PENSE QUE L'UN DES DEFIS RESIDE DANS LE FAIT QUE LES GENS ABSORBENT DES DONNEES ET DES INFORMATIONS, MAIS QU'ILS NE SE SOUVIENNENT PAS TOUJOURS CONSCIEMMENT DE CE QU'ILS ONT UTILISE A UN MOMENT DONNE ET DE LA MANIERE DONT CELA A INFLUENCE LEUR ACTION OU LEUR DECISION. – CHRISTINA WILLE, INSECURITY INSIGHT

Participants of our June 2024 francophone exchange-day testified that it still often occurs in humanitarian operations that the decision is made and then the justification found with the data[15]. Few actors have the time and resources to properly analyse data: the organization specialized in data quality & accountability Ground Trush Solutions gave the example of a perception survey on cash distribution, in a specific area in Nigeria, where surveyed people expressed a preference for vouchers. Due to very wide acceptance of cash distribution as a reference approach, response actors used this information as a starting point for a discussion, which proved a smart decision as later qualitative research showed a more nuanced picture – people didn't like the *specifics* of cash distribution in the

[14] « Confirmation Bias Is The Enemy Of Exploratory Data Analysis », Jonathan Davis, Towards Data Science, Septembre 2020
[15] Program data management seminar, June 2024

area by NGOs, not the principle. The solution was rather to improve cash distribution rather than going back to vouchers… But, would actors have taken the same care and time to verify this information if not "attacking" such a well-recognized approach as cash distribution? This links to psychological mechanisms such as confirmation bias that are also being analysed by specialists[16].

Another way to assess this question is to ask the data specialists themselves: during a workshop organized with data management specialized organizations from the H2H Network, on a panel of 18 experts from 10 organizations, they graded 2.9 out of 5 their capacity to evaluate the impact of the data/tools/support they provide (with a significant part at very low scores):

> I THINK ONE OF THE CHALLENGES IS THAT PEOPLE ABSORB DATA / INFORMATION BUT THEY MAY NOT CONSCIOUSLY REMEMBER WHAT EXACTLY THEY USED WHEN AND HOW IT INFLUENCED WHAT THEY ENDED UP DOING OR DECIDING. – CHRISTINA WILLE, INSECURITY INSIGHT

Most data specialists indeed struggle to evaluate the impact of their action. For instance another data-specialized H2H organization hired an external evaluator to try to measure their impact better but only came out with limited results; at CartONG we also very seldom have the capacity to do proper impact evaluation of our own projects (most of the time metrics are based on products realized or people integrated in the process, but not on impact of said products or long-term impact on these people), which is obviously frustrating for data-driven organizations like us… CartONG and MapAction have developed a generic methodology on how to evaluate data components of programmes[17], but it's not fully tackling yet the question of impact. The capacity to assess the impact of data also varies depending on contexts, it is very clear on acute emergency with an initial lack of data (e.g. pandemic responses like Ebola), and much less visible for wide-scale transversal crisis.

The specialized organization The Engine Room concurs on the lack of evaluation and accountability for implementing data and technology solutions, also pointing

[16] "Confirmation Bias Is The Enemy Of Exploratory Data Analysis", Jonathan Davis, Towards Data Science, September 2020
[17] "Guide to evaluating Information Management components of programmes", CartONG and MapAction, 2022

the associated risks: “each wave of shiny new technology brings additional cost and complexity – more efficiency or improved outcomes are never guaranteed”[18]. Too often are humanitarian organizations rather running from one tool to another in their quest for efficiency, without ever having the time and resources to assess the impact, sustainability, constraints, risks (including for fundamental rights), and integration with existing efforts.

In fact, very few long-term evaluations exist concerning the transformative impact of data technologies for the humanitarian sector. An ALNAP synthesis on a sample of 540 humanitarian innovations found out that evidence of impact was available for only 16% of them, and that funders usually had no information on the outcomes of innovations after the end of their grants[19]. If we check for instance what was listed five years ago as the top innovation trends for future years[20], while some are still seen as foundational (internet access, improved data analytics), one could question if others have really transformed that much the work of most humanitarians in the planet (for instance virtual reality, 3D printing, or even drones), except for very specific sectors or purposes.

Focus: cash distribution, embracing technology for a specific purpose
The impact of another “big trend” of the past 5 years is interesting to explore: blockchain. While we’re still far from the widespread adoption some predicted in the 2010s, due notably to the very high-tech expertise required (i.e. highly technical specialist with payroll CSOs rarely can afford), blockchain has found its place in several specific sectors, in particular for cash distribution[21]. We can use cash transfer as an example to highlight that having a very clear use case justifying a technology is a key factor to ensure its relevance and impact. Without going into all the advantages (speed and scale of response in particular) and disadvantages (potential data protection risks, inflation, targeting, etc.) of digital payments, it is a technology that is technically very complex, and requiring the collaboration of many types of actors to be deployed (donors, private service providers, CSOs, etc.). It also symbolises

[18] "In the humanitarian sector’s search for efficiency, are we falling short?", Laura Guzman, The Engine Room, May 2023
[19] “Assessing the promise of innovation for improving humanitarian performance: A 10-year review for the State of the Humanitarian System report”, ALNAP/ODI, October 2023
[20] “7 tech trends that are transforming humanitarian aid”, MercyCorps Blog, January 2019
[21] “The State of the World’s Cash 2023 – Chapter 7 Data and digitalization”, CALP Network, November 2023

well current difficulties that can occur with new technologies (risks of mosaicking of data, cybersecurity stakes, etc.).
Its humanitarian ecosystem of actors has however – compared to other technologies – invested strongly on risk analysis and associated learnings to ensure it is adopted and implemented in the right way. The associated CALP network offers a wealth of resources (such as the very practical Data Responsibility CVA toolkit[22]) and reflexions on the topic that have been a driving force for sector discussions on many other technologies and approaches used.
That being said, it is worth noting that even this success story has not yet scaled to the level it could, as cash distribution is estimated to be utilized only at 50% of its potential, "not achieving its full potential because the system is slow at scaling effective ways of working"[23].

A significant example of more mitigated results is one of the most emblematic – and controversial – technology adopted in the humanitarian sector in the past years: biometrics. According to the 2023 report of the Engine Room[24], "much of the motivation for the use of biometrics rests on claims that it will aid in de-duplication efforts, fraud control and anti-corruption, with limited evidence demonstrating a clear, positive benefit in these areas". Beyond this, as their research shows, "the adoption of biometric technology continues to put data subjects at the greatest risk", with questions such as decision-makers' technical literacy, coherence of organisational policies, patchy implementation and inadequate funding by donors for safety measures as blocking points.

Generally speaking, as we'll develop later (cf. chapter 9.12) on our relationship with digital technologies, it is not the *nature* of each technology that leads to its adoption and eventual impact, but the innovation adoption *process* around it, whether it provides the necessary conditions (in terms of investment, lasting partnerships, sufficient testing, etc.) or not.

[22] "Data Responsibility Toolkit: A Guide for CVA Practitioners", Linda Raftree, Anna Kondakhchyan, CALP Network, March 2021
[23] "Assessing the promise of innovation for improving humanitarian performance: A 10-year review for the State of the Humanitarian System report", op. cit.
[24] "Biometrics in the humanitarian sector"; Quito Tsui, The Engine Room, 2023; Check also the summary video here: "Biometrics in the humanitarian sector", The Engine Room, 2024

5.2. DATA QUALITY: AN EVERLASTING QUEST?

i. Biases in data and underrepresentation

The question of the quality of data could seem like a cliché by now, or as a problem from the past our industry has solved. However, and even before entering into technical questions, new perspectives on humanitarian action have of course shown that there might be more issues than initially perceived with what we can consider "quality data".

The first issue that has emerged recently is the lack of local definition of the data production agenda in the humanitarian sector. As explored in CartONG's previous research "Changing the outlook: for a local approach to data"[25], data production is often driven more by accountability to international actors rather than use by local actors, whether they are CSOs or governments. This can lead to distortion in the very perception of the humanitarian sector, as illustrated in a case study by Development Initiatives on International Aid Transparency Initiative and OCHA's 3W (Who/What/Where) data in Somalia[26]. Local actors were figuring less in 3Ws than their actual field presence, and disappeared altogether from IATI global reporting because they didn't appear in the financial data on which the latter based…

Another major bias that can skew the entire data quality chain in the humanitarian sector is of course language.

A case study of the distorting impact of language: TWB in Nigeria
Translators Without Borders/Clear Global has provided numerous examples of humanitarian projects whose accountability, relevance, or simply feasibility are challenged by the lack of proper translation in language understandable to their beneficiaries[27]. For instance in 2017 in Nigeria, Translators Without Borders verified the comprehension by local enumerators of a survey's questions for enumerators speaking English and Hausa, and their capacity to translate it into the 28 local languages, showing a lot of inaccuracies or

[25] "Changing the outlook: for a local approach to data", op. cit.
[26] "Improving the visibility of local and national actors in humanitarian aid data", Development initiatives, July 2021
[27] "Listen and learn: The link between language and accountability for the future of the Grand Bargain", Translators Without Borders, June 2021

differences in translation (including concepts that don't even exist in other languages such as accountability, or technical terms)[28]. On top of this, enumerators often had to rely on a third-party to translate in a language they didn't master (with again, no quality control possible on the quality of the translation provided by this person selected on the go), or even to skip some interviewees when not finding a translation proxy.
All of this distorts the data produced, with some populations structurally under-represented in the data collected (women, less educated populations, minorities…) and a general question on the quality of said data when such differences can be seen in the way it is collected. These questions on the actual validity of the data on which the sector is building his assessments and evaluating itself can of course be extended to many crises, in particular those in regions with high linguistic diversity. It is however difficult to make a comprehensive assessment, since even having a global database of what languages are spoken where it still a work in progress…

> THERE ARE SYSTEMIC WEAKNESSES IN OUR SECTOR'S APPROACH TO ACCOUNTABILITY, RENDERING WHOLE GROUPS VOICELESS WITHIN THE SYSTEM, SIMPLY BECAUSE THEIR VOICES ARE IN ANOTHER LANGUAGE. – CLEAR GLOBAL

Another key factor that can limit accuracy of data and that is still under-measured, even if the situation is progressing, is of course gender. As measured in our 2020 study, women are under-represented in the data management sector, as in many so-called "technical" professions. This influences the way tools and methods are built, for instance with AI. The data that feeds large language models systems such as ChatGPT are indeed built on a corpus that reflects the inequalities of Western societies, including on gender[29].

A good example of these possible biases are in the impact of data in the allocation of aid. For instance in climate action, a research by Development Initiatives has shown that not only did the *gender mark* used for official monitoring of aid not

[28] "The power of speech", TWB's response in Nigeria, Translators Without Borders
[29] "How can we apply feminist frameworks to AI governance?", Linda Raftree, MERL Tech, September 2023

correctly track the projects in question, but it failed to measure the needs and lived experiences of women on the frontline of climate change. The authors advocate in favour of the importance of an intersectional approach to data – taking into account the combination of different types of discriminations and dominations (citing the example of the higher rate of child marriage in areas prone to recurrent droughts)[30].

> WE HAVE SEEN FIRST-HAND THE VALUE THAT THE INCLUSION OF VOICES CAN HAVE ACROSS THE DEVELOPMENT DATA VALUE CHAIN – FROM CONCEPTUAL DEVELOPMENT TO COLLECTION, AND FROM ANALYSIS TO USE – AS WELL AS THE VALUE OF NOVEL SOURCES OF DATA DIRECTLY PROVIDED BY CITIZENS AND COMMUNITIES IN THE FORM OF CITIZEN-GENERATED DATA. - DEVELOPMENT INITIATIVES

To conclude, as reflected recently by the Inclusive Data Charter's members when reviewing 5 years of impact, working on inclusive data means going beyond simple data disaggregation, to work on the challenges of intersectionality of exclusions (examples of language + gender + racism evoked earlier), and more broadly on the challenge of digital inclusion[31].

ii. Data quality down to the last mile

Notwithstanding these "new" quality questions linked to data sources, more "traditional" data quality challenges remain: faking GPS points, inventing an interview, etc. "Data cooking" still exists and can have a significant impact on results, even though it has diminished with digitization of processes[32]. Most of the frameworks on data quality focus on methodological aspects: relevance for the sector, coherence, timelessness, punctuality, accessibility, interoperability, etc. The reliability / credibility of sources remains a more debated concept, if extending

[30] "When the data doesn't tell the full story: improving gender-responsive climate finance", Mariam Ibrahim, Fionna Smyth, Claudia Wells, Euan Ritchie, Development Initiatives Blog, November 2023

[31] "Reflecting on five years of the Inclusive Data Charter", Tichafara Chisaka, Global Partnership for Sustainable Development Data, March 2024

[32] "Who owns data in Somalia? Ending the country's privatised knowledge economy", Mahad Wasuge, Ahmed M. Musa & Tobias Hagmann, Somali Public Agenda, July 2021

only to the organization or to the data itself (and thus the way it is collected). However, a data quality specialist testified that when asking in the field, in most contexts local data enumerators declared that before working for them, no international organization proceeded to quality control their data collection work in any way…

OCHA's Centre for Humanitarian Data has come out with several recommendations to strive for data quality, in particular through the leveraging of existing quality measures and domain expertise, but also automation of certain tasks (e.g. pre-filling or requiring certain metadata) and displaying in simpler way data quality to help select the most qualitative[33]. And multiple projects exist to strengthen the quality of base data in everyday use in the sector, such as MapAction's and University of Georgia's work on Common Operational Data-Administrative Boundaries[34].

However, the incentive structure of the humanitarian system, which strives always to increase efficiency by streamlining processes, can have negative side effects on data quality. For instance, the strong push for collaboration to reduce redundancies is of course positive, but has the side effect of reducing the sources of data on an issue, which can be a problem as it doesn't allow comparison of sources and crossing of information. Specialists thus now have to push donors to not reduce data collection too much to keep multiple data sources: "data cannot be treated as a commodity, something that would be just handed out to an organization without context" (Ground Truth Solutions).

The same organization is not very optimistic on the capacity of the sector to advance on fixing these structural issues. Data quality is seen by the general opinion of the sector as a "problem fixed" with digitalization (or big data, or remote sensing, etc.), and it is quite difficult to get traction or funding for efforts that question this assumption. Even though several data specialists from CSOs testified that they still struggle to get buy-in from Country Directors… who are paradoxically in the front row to witness the frequent unreliability of data!

[33] "Quality Measures for Humanitarian Data", Data Nutrition Project, Centre for Humanitarian Data, September 2023
[34] "Strengthening data quality for shared humanitarian data sets can reduce human suffering", MapAction, July 2024

So, in conclusion, the level of data quality pursued (and resources dedicated to it) must always be proportional for the purpose targeted. As detailed during webinars CartONG conducted to train CSOs on data management[35], and in particular data collection, the minimal quality threshold should always be defined based on the purpose for which the data is collected (reporting, monitoring & evaluation, piloting, communication, research, etc.).

Data collection should not be seen as the end goal by CSOs, either because they are pressured into doing it, or to appear competitive in the innovation front. To take an example, while mobile data collection has allowed a huge leap forward in data quality with its embedded quality-oriented features, there is no evidence it has allowed clear improvements in other steps of the data management process (analysis, identifying biases, etc.). Another example is the development of remote data collection/monitoring systems, sometimes linked with predictive models, at some point seen as the silver bullet to reach hard to access areas. Experience has shown that they will not be able to replace field/local data collection, the lack of granularity not allowing some purposes required in many situations (including obviously monitoring / more programmatic aspects!). Hence the need for reinforced data literacy efforts (cf. chapter 9.5) to promote better use of data, but also to always question, “why do we need this data?”.

iii. Où sont les données qualitatives et secondaires ?

Another gap on the question of data quality is the common underuse of qualitative data and secondary data. Quantitative data is often seen as simpler to collect – requiring “only” technical skills –and is therefore largely favoured by humanitarian actors.

This is also linked to a larger bias already mentioned to also focus on data that can be used to measure cost-efficiency, leading to a general “obsession on counting things” as a Ground Truth Solutions puts it. The higher actors are in the decision-making system (e.g. donors, UN agencies, leaderships in general), the more they are under this accountability pressure that leads to a simplistic vision of monitoring & evaluation, and of the data behind it, pursuing only quantitative approaches. Methodologies such as logical frameworks push for pre-defined

[35] [In French] "Webinaire - Trucs et astuces pour améliorer votre collecte de données sur mobile (focus KoboToolbox)", CartONG September 2023

indicators and quantitative data, leaving little room for contextual analysis. This can also cause a challenge for smaller, local CSOs who lack data collection capacity but often have an excellent contextual knowledge of their environment: international CSOs and donors seldom adapt to this. Methodologies focused on measuring change and quality are still only emerging in the sector (cf. also chapter 9.2). In fact, more qualitative approaches are mentioned by CSOs we interviewed as a solution to reduce the gap between fields and headquarters, and between organizations and populations – including sometimes by reducing the digitalization of processes for a more human touch.

Nonetheless, Ground Truth Solutions does not witness resistance to qualitative approaches on principle: when well designed and introduced to the user, qualitative data can go quite deep and make important differences in actors' perceptions. At the end of the day, the capacity to defend the methodology, and even to prepare decision-makers to the data they will receive, might be as important as the data itself… Specialists recommend that beyond methodologies, humanitarian actors work on their posture to generate spaces of discussion and encourage proactive involvement of the community[36].

> WHILE THERE IS SOMETIMES A TENDENCY TO SEE QUANTITATIVE APPROACHES TO EVIDENCE GENERATION AS 'HARD' AND QUALITATIVE APPROACHES AS SOMEHOW 'SOFT', THE USE OF QUALITATIVE APPROACHES SHOULD NOT BE AN EXCUSE FOR LACK OF RIGOUR. UNFORTUNATELY, MANY HUMANITARIAN EVALUATIONS DO NOT USE ACADEMICALLY RECOGNIZED QUALITATIVE METHODS, AND FAIL TO MEET BASIC QUALITY STANDARDS RELATED TO ACCURACY, REPRESENTATIVENESS, AND RELEVANCE. – ALNAP[37]

The same can be said about the use of secondary data (i.e. data preexisting to research, e.g. collected for another project), which is still lagging in the

[36] "Rediscovering the power of words to generate knowledge - Qualitative methods, robustness and accountability - towards a paradigm shift in MEAL", Sophie Mareschal (Terre des Hommes), IM Portal Blog, July 2023

[37] "Strengthening the quality of evidence in humanitarian evaluations", Ian Christoplos, Paul Knox Clarke, John Cosgrave, Francesca Bonino, Jessica Alexander, ALNAP, May 2017

humanitarian sector. This is due to technical sharing and legal issues (cf. chapter 9.8) but also because quality secondary data review requires significant resources for data mining, cleaning, assessment, consolidation and finally analysis[38].

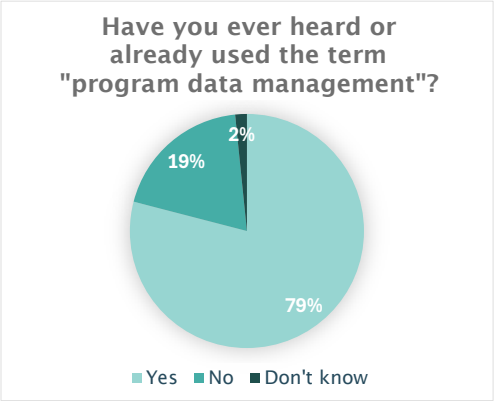
AI-supported tools seem a promising field to reduce this barrier and facilitate access to secondary data, with projects such as Deep.io[39] (consortium of several major UN agencies, IFRC, and service providers, created in 2022), GANNET[40] (Data Friendly Space, 2023, under development) or SOPHIA[41] (ACAPS, 2024). While all these platforms show strong potential and are already starting to gain users and complete their database (more than 90,000 sources and 6000 users for Deep.io at the time of writing of this study), it is still too early to discuss their actual impact, and no case studies or evaluation have been produced on them yet.

[38] Cf. for instance the methodology by UNHCR: "How to conduct a secondary data review", UNHCR Assessment and Monitoring Resource Centre, November 2023
[39] "DEEP – a collaborative analysis platform for effective aid responses", DEEP, 2022
[40] "GANNET – AI for humanitarians, by humanitarians", DataFriendlySpace
[41] "Welcome to SOPHIA", ACAPS

6. A FIELD UNDER PROFESSIONALIZATION

6.1. "PROGRAM DATA", A BETTER UNDERSTOOD AND RECOGNIZED CONCEPT

One of the stake of our 2020 study was to clarify the terminology of our field of work, which we did, in particular around the concept of "program data". The aim was to make the scope more specific than the very wide "information management" (IM) term that can englobe much more than program-related data (financial, administrative, logistical, HR data, etc.) and less framed as "humanitarian" than IM.



The concept seems to have filled a need in the meantime, with 88% of the actors who answered our survey working in humanitarian responses using it (but in fact only 67% working in international development); and 100% of humanitarians understanding it (91% of development professionals). To be noted though that since the survey was disseminated primarily through CartONG's networks, there is certainly a bias with many respondents, in particular the humanitarian ones, more exposed to our previous research work.

We can however mention that international development CSOs still lag behind in terms of endorsing the concept, using a wider variety of terms (such as "data", "project data"…). This can be related to the lower structuring of the data-related professional field in these organizations – while humanitarian familiar with the twin concept of "information management" can more easily grasp it. Different interviews highlighted indeed that the question of the terminology is often linked to the level of onboarding of the organization on the topic – as soon as it starts clarifying where data stands in terms of roles or when a focal point starts being appointed it leads to a clearer endorsement of vocabulary. There are therefore still efforts to make to have a more common vision between the actors (in a nexus vision) on this data aspect – even though CartONG's work was in part designed to bridge this gap.

Most organizations clearly connect the concept to the 2 sides of the definition we established in our previous study: (i) based on a multi-step data management

cycle and (ii) aimed at improving decision-making and the quality of humanitarian actions. Even smaller and local organizations link it clearly to their monitoring & evaluation, assessments, programming, and also accountability to the populations and to their donors.

In terms of regional specificities, it is interesting to notice that in the francophone sphere, where the lack of structuring of the professional field (and its vocabulary) still can be seen (more than two-thirds of the respondents from our survey who didn't know the term were francophone), the gap has significantly reduced – hopefully the impact of the more understandable "data management" concept. According to a network head, this persistent lack of adoption of the concept in the francophone sphere (also in part because CartONG is "taking the space"!) is more a consequence of a general lack of resources in a world of competing priorities, rather than a lack of understanding and interest on the topic as was the case before.

Expert support actors also seem to approve the centrality of the concept: (program) data management, and the linked concept of Information Management is also the way they spontaneously identify their work (workshop with H2H data specialized organizations). They open the definition to connecting topics such as knowledge management, data analysis, data science, data monitoring, data responsibility, GIS… reflecting different professional approaches depending on their specialty/training, all linking though to the same core idea.

6.2. GROWING SKILLS

To explore the questions of tools and methods that are currently seen as part of data management, we will follow the different steps of the data management cycle, as we summarized them in our 2020 study, focusing on the technical steps:

Data management cycle

Data collection

Main pattern
Direct surveying by CSOs staff – whether through paper, mobile or computer – remains by far the most used method.

Local CSOs specificities
Still using more paper forms.

International CSOs specificities
Using more mobiles, and up to 2× as much "advanced" data acquisition methods (complaint mechanisms, chatbots, remote sensing).

Analyze data

Main pattern
Tools with a more advanced visualization component (GIS, cf. the focus below, or statistics/dedicated tools) are progressively being adopted; an important diversity remains depending on the sector of intervention of the organizations.

International CSOs specificities
Using more "technically intensive" tools in terms of resources (skills, dedicated staff & infrastructure, license costs) and procedures / integration of databases (linking with compliance to donor or authorities' requests, including feedback & complaint mechanisms).

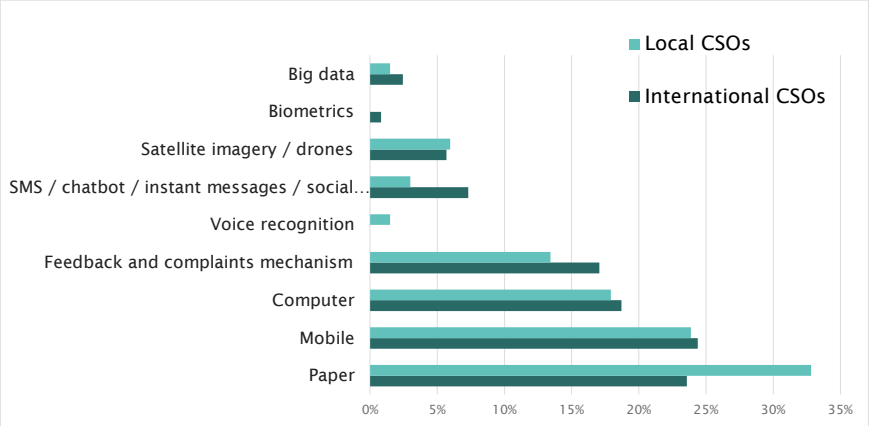
Data processing and storing

Main pattern
Still mostly done using basic offline tools such as Excel.

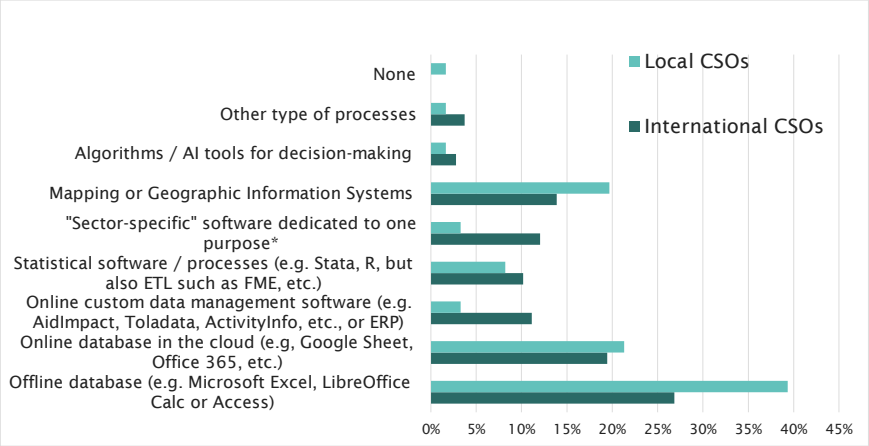
Local CSOs specificities
Particularly using Excel.

International CSOs specificities
Using relatively more online database management and "advanced" processing methods (aggregation tools).

Looking at the numbers, here are the share of respondents that identified a regular use of the following data collection processes in their organization:



And similarly for data processing & analysis processes:



Focus: GIS as a professional sub-field within data management

The greatest increase between 2020 and 2024 that we can note through our survey in terms of usage is on Geographic Information Systems, with 44% of our respondents using it in a least one project (compared to 33% in 2020), in particular local CSOs. It is interesting to note that for local CSOs, it is the only tool they use more than international ones apart from Excel, which shows the

importance of such a tool for adequate local contextual understanding (and can raise questions as to why international CSOs are not using it more for such purposes!).
This in any case corroborates observations by several actors interviewed for our study, but also CartONG's own track record of trainings requests: more and more organizations (local or international) are looking to develop basic GIS skills, sometimes for a large number of their staff. We can mention here the very interesting cross-donor MAPME initiative[42] (by KFW, GIZ, AFD and IRD) to provide free and open access GIS and earth observation tools to international development cooperation actors, in particular for evaluation, that shows how such actors are convinced of its added value as well as of a mutualised approach to the topic. To get back to CSOs, their needs can sometimes concern very basic geographic data tools such as Google Earth/Maps/GoGoCarto, but even professional GIS tools as QGIS and ArcGIS have sufficiently simplified their interfaces to allow a certain level of usage by non-specialists.
This expansion is probably linked to the fact maps are an easy-to-grasp decision-making tool that helps present different types of information in one place with a new perspective, which helps justify the needs for training inside organizations. GIS remains however a separate profession, both in terms of education and then careers paths. This can lead to relatively autonomous dynamics from the rest of the data ecosystem, including within an organization (cf. example of Médecins Sans Frontières, ICRC or various UN organisations to mention the most structured actors of the sector on the topic) or within a sector (cf. for instance mine action within a CSO like Humanity & Inclusion). In smaller organizations, the development of GIS or mapping also tends to be more dependent on individual profiles interested in it.

It is interesting to note that the "revolutionary" new technologies that we heard about the most in the past years (big data, remote sensing, biometrics, voice recognition, algorithms, AI or chatbots…) remain low or even anecdotal in terms of use for CSOs, be they local or international. It however doesn't mean that CSOs are not at all interested in using the information and knowledge that can be made available (when that is their vocation), just that this type of technology clearly

[42] "MAPME - Maps for planning, monitoring and evaluation in development cooperation", MAPME

remains more the prerogative of big International Organisations or specialized support ones, as the level of skills, resources and systems required is at a whole different level.

Regarding the comparison of skills vs. needs, it seems to have improved significantly since 2020: while in our previous survey, less than a quarter of the respondents evaluated these skills as sufficient, this figure now raises to 62% (for HQ staff) and 53% (field staff). Only 35% (at HQ) and 47% (field) of staffs in charge of data are estimated to be only partially qualified. This can be explained by the fact that all CSOs are using technology more widely than a few years ago, even though it is for the most part very easy and accessible tools, and that they are therefore more widely part of the skillset of CSOs.

> In recent years [our NGO] has invested considerable resources in data management, and we've seen a considerable step up in skills and knowledge both at head office and in the field. Although everything is still far from perfect, I think that we are relatively well equipped to face up to the various challenges in this area, and to continue to make progress in this field. – HQ Data focal point from a humanitarian CSO

In recent years [our NGO] has invested considerable resources in data management, and we've seen a considerable step up in skills and knowledge both at head office and in the field. Although everything is still far from perfect, I think that we are relatively well equipped to face up to the various challenges in this area, and to continue to make progress in this field. – HQ Data focal point from a humanitarian CSO

The gap between HQ and field staffs can be explained by a difference in specialties: HQ data staff are often linked to dedicated functions (M&E, IT, data protection…) with a higher specialization, while field data staff are often mixed positions including M&E and some IT or program component, which doesn't allow as much specialization.

The gaps are however lower in some countries or regions, such as the Midde-East, where the local economic and social development and associated level of digital expertise, the number of local staffs hired by international CSOs and other agencies, and the turnover within the sector, means that a high number of qualified

technicians are available, and move skills, tools and ideas from one organization to another. This also reflects in the funding of the innovation ecosystem: in a sample of 540 innovation projects reviewed by ALNAP over a decade, 36% were located in East Africa, 22% in South and Southeast Asia, and 10% in the Middle East (and only 6% in the rest of Africa for instance)[43]. Relatively easier operational contexts such as Uganda (16%), Kenya (7%), the Philippines (10%) and Jordan (4%) were highly represented, while only 6% of the projects were tested in level-3 emergencies.

Challenges however remains, similar to 2020: the most mentioned one being on methodologies (46%), the technical and strategic aspects appearing far behind at respectively 18% and 24% of respondents (while 12% of respondents cannot identify precisely the causes of these difficulties).

This is consistent with the learnings issued from CartONG's dedicated work on program data Human Resources in 2021[44]: while most organizations identified technical challenges and solutions as their priority, for data specialists it was indeed these methodological and organizational (including HR/hiring) aspects that were a priority. And while IM is now well identified as a dedicated function in the humanitarian sector, it still lacks the dedicated initial and vocational training (that other fields such as M&E have), and the bridges between the humanitarian sector and others (in particular the private one) are still complex. Even though the appearance of positions framed around data in large CSOs (data scientist, data protection officer, data analysts, etc.) can help these transitions.

6.3. THE ORGANIZATIONAL LEVEL'S CHALLENGE

Individual skills can only go so far without an uptake of the data question at the organization's level. At the heart of this is the question of who owns the responsibility of data management in organizations. Important differences exists depending on the organizations: while a significant share of our survey's respondents now have structured a team of data specialists in the field (37%), either through a team (34%) or through an individual (22%), on the other hand there are still a significant number of organizations where data management has

[43] "Assessing the promise of innovation for improving humanitarian performance: A 10-year review for the State of the Humanitarian System report", op. cit.
[44] "HR pack – Program data management for humanitarian aid and international development CSOs", CartONG, March 2021

no defined person in charge (15%), or else work with volunteers/externalized support (respectively 5 and 2%). Again, and logically, the professionalization of the function is logically much higher in international CSOs (41% having a team in the field and 38% at HQ) than in local CSOs (25% for both), who in turns rely more on volunteers for it (17% vs. 0% for international CSOs).

The type of profiles in charge of data management at CSO HQs are a good indicator of both this growing specialization, and the HQ/field gap: the most common occurrences are positions dedicated to data (45%) or to M&E (52%); followed by positions dedicated to innovation/digital in general (24%) or IT (28%). Positions linked to programs (17%) and reporting (14%) come next. On the contrary, at field level 93% are linked to M&E and 20% to programs, with much lower occurrences of other profiles such as innovation and IT (20% each), showing a lower level of specialization.

These differences can have significant consequences: for smaller and local CSOs, new requirements such as cybersecurity, advanced analysis or consistent methodologies are too complex for volunteers or program staff with no dedicated training or support. Funding remains a key gap for them, as they usually don't manage to integrate data support or tools when building their budgets.

Finally, one last interesting angle to analyse organizational endorsement of data management is to observe the potential variations within the CSOs. According to our survey respondents, a majority (54%) of organizations live differences in appropriation of data management depending on the sectors of intervention, while 39% have consistent appropriation through the whole organization. While we didn't ask the same question in 2020 and cannot compare directly, the rate of CSOs with consistent practices organization-wide thus seem to increase, likely linked with the growth of organizational M&E standards and functions.

As had been observed in 2020, some fields however remain more "data compatible with tools and standard practices available (e.g. WASH, demining, distribution, cash transfer, etc.) while sectors linked to community activities, social services or education, which are harder to measure and require a more qualitative approach, still have less data use. This can also be linked with the differentiated importance of a given sector within each organization, the "core work" of each CSO often being more in need of data for accountability and advocacy – we can quote the example of a WASH specialized who created a dedicated position on MEAL and quality for the WASH department.

Similarly, various organizations also mentioned persistent internal differences between their country missions, depending on their size, resources, structuration and the interest of local leadership (head of mission). Location also has an impact with the influence of authorities (e.g. when requiring the use of standardized data collection/aggregation tools such as DHIS2), or the availability of local skills (cf. previous chapter). And the same can of course be said for international CSOs with different sections/headquarters.

> LARGE DE-CENTRALISED ORGANISATION STRUCTURE BRING CHALLENGES TO THE DEGREE OF COMPLETENESS AND CONSISTENCY OF DATA ACROSS THE ORGANISATION.
> – WELTHUNGERHILFE

In general, the different testimonies we received through the survey and our interviews show that the situations remain very diverse. For instance, this big international CSO created around the food distribution field later diversified: while their mental health sector had formalized standardized data collection models but still used different tools, their WASH sector uses a standardized tool (mWater) but no standards in data models; and their "food security" and "nutrition & health" sectors still didn't have either models or tools standardized.

As we can see, the continuous movement towards more structuring of the data function, definition of policies and strategies and dedicated positions, touches a larger panel of CSOs, including smaller / medium ones. But as various interviewees from different structures testified, this systemization, while now understood as necessary in many organizations, remains a work in progress. As dedicated positions (or positions with a strong data component such as M&E) grow, new questions arise such as the coordination between HQ and field data (harmonization of practices), siloes appearing (e.g. inside M&E departments), disconnection between evaluation and programs (monitoring), differences between sectors in an organization, etc.

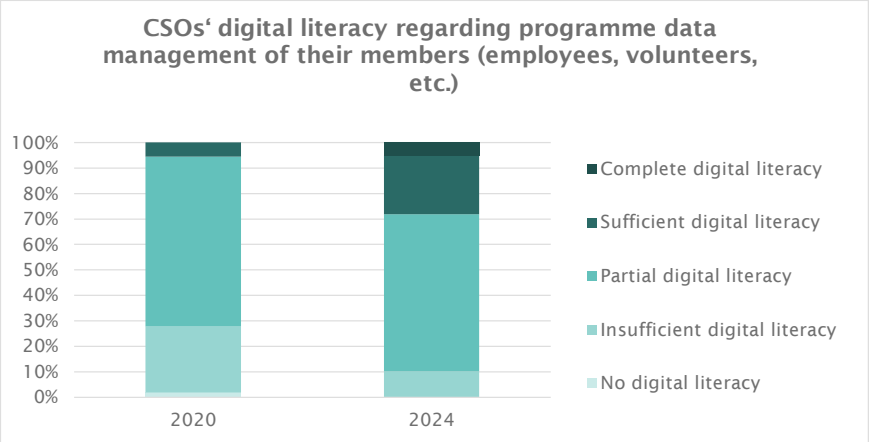
Finally, the question of organizational endorsement of data is of course necessarily connected to its ownership by the organization's leadership. This aspect will be covered in detail in chapter 9.4 on CSOs' strategies. Indeed, the more complex and crosscutting challenges we'll cover in chapter 9 often require long-term investment in human or policy resources. It is therefore critical that leadership defines a consistent strategy that builds on the essential blocks of data literacy (cf. chapter 9.5) and sustainable systems (cf. chapter 9.11), rather than pursuing

short-term tech challenges that are “fancy” in the sector, as data specialists have often witnessed in the past.

6.4. A PROFESSIONALIZATION CONNECTED TO DATA LITERACY

The professionalization of the humanitarian sector in terms of data primarily occurs through data specialists. However, they cannot uplift the whole sector alone: raising the skills of the sector depends on a more global approach towards data literacy. Data literacy consists of “the ability to read, work with and analyse, argue with data”[45]: without basic data literacy spread at the organizational level, the work of specialists is pointless.

Our survey respondents – who were in majority data specialists within CSOs, or at least “data-curious” – consider there is still much work to be done in that direction. Indeed, for more than two-thirds of them, their colleagues had a partial or insufficient digital literacy regarding data.



While these figures have improved since 2020, they show that many organizations still don’t have an organization-wide approach towards data literacy, relying usually a lot on “evangelization” from a handful of passionate profiles. One of the few sector-wide survey conducted on the topic in 2019 by OCHA’s Centre for

[45] “Why data literacy is important in the aid sector”, Sylvia Musula, IM Portal Blog, CartONG, July 2021

Humanitarian data lead to similar conclusions, with a very common use of data by most humanitarian professionals (98% of respondents using data all the time or occasionally), but between 30 and 50% of most profiles also reporting challenges in the common dimensions of data collection, analysis and data quality assessment/improvement[46].

> SOME PEOPLE IN MY ORGANIZATION ARE "VERY CONVINCED" AND HAVE SOLID EXPERTISE, BUT THEY ARE NOT IN A POSITION TO MAKE DECISIONS AND HAVE DIFFICULTY CONVINCING DECISION-MAKERS, EVEN THOUGH THIS HAS LARGELY IMPROVED OVER THE LAST TWO YEARS. NEVERTHELESS, THE CHRONIC WORK OVERLOAD OF THESE PEOPLE MEANS THAT THEY DON'T HAVE THE TIME TO "SIT BACK" AND THINK ABOUT FUNDAMENTAL ISSUES (STRATEGY, HIRING NEW PEOPLE, ETC.). – A DATA SPECIALIST FROM AN INTERNATIONAL DEVELOPMENT CSO

It is also interesting to note that no significant differences between international and local CSOs can be seen on this question: this is a sector-wide challenge, connecting to various dimensions in terms of leadership, resources, policies, etc. The gap seems more important between francophone and anglophone CSOs, the former being traditionally less exposed to a "data-driven" culture and thus less pushed for a wider appropriation of data literacy. CartONG identified in a previous research several causes for this persistent gap: "the greater prevalence of ICTs in English-speaking societies, larger employment and skill pools in the field of data, a larger bibliography available in the English language, etc."[47]. We will go into more details in chapter 9.5 on risk and opportunities associated with data literacy.

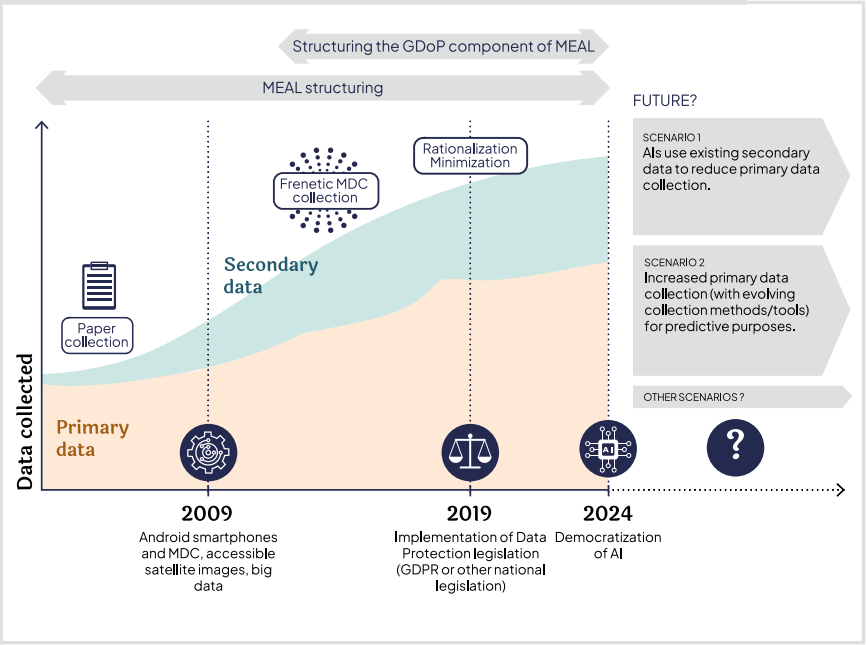
To summarise the past couple of decades on the topic in terms of structuration path that many CSOs are following at their own pace, we have worked on this illustration, that looks into some of the key tech appearances (Mobile Data

[46] "We are all Data People: Insights from the Data Literacy Survey", Katelyn Rogers, Centre for Humanitarian Data OCHA, April 2019

[47] "HR pack – Program data management for humanitarian aid and international development CSOS – Framing the key issues and getting familiar with the toolbox", CartONG, March 2021

Collection, Business Intelligence, more accessible Satellite imagery and Geographical Information technologies…) and events (data protection legislation such as GDPR) and their impact on the professional field, in terms of type and quantity of data sources, as well as working methods. The big question today being of course how Artificial Intelligence will change things – on many topics, but in particular in terms of working methods related to the data cycle, be it in terms of quantity and quality of data available but also how data will be analysed.

The history of data management through the prism of data quantity



7. NEEDS: SOLUTIONS & RESOURCES COMPATIBLE WITH SECTOR SPECIFICITIES

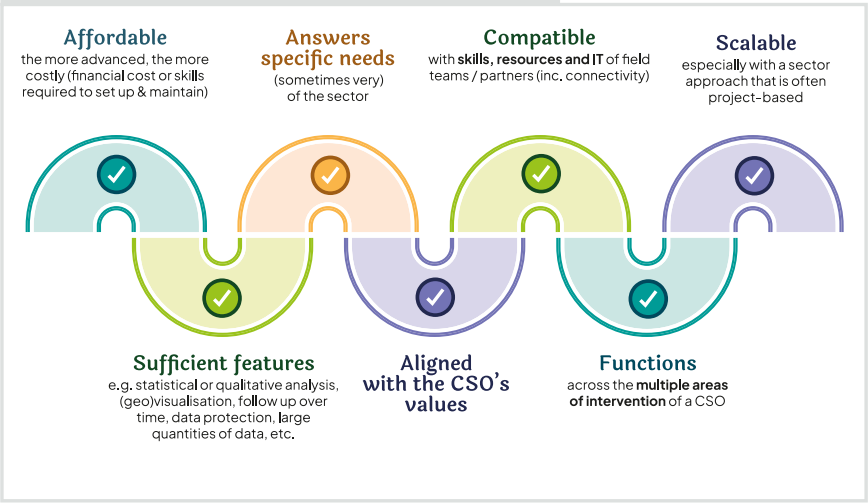
7.1. ENSURING TECHNICAL SOLUTIONS...

The relative profusion of technical solutions in data management developed over the past 10 to 20 years could let us think there is no challenge anymore for CSOs to fill their technology needs. However, more than 50% of our survey respondents have mentioned having "at least sometimes" not found a solution to their needs, a figure that goes up to 66% for local organizations.

> THERE ARE INSUFFICIENT TECHNICAL AND FINANCIAL RESOURCES FOR THE IMPLEMENTATION OF SOLUTIONS, AS THESE ARE OFTEN NOT INCLUDED IN DONOR FUNDING OR SUPPORT – RESPONDENT FROM A LOCAL CSO

This does not necessarily mean that the solution doesn't exist. In fact, according to our respondents, a multiplicity of requirements can complexify the search for an appropriate solution. Tech solutions indeed need to cover the following requirements:

Requirements for an appropriate technical solution



These multiple requirements also mean CSOs need to have transparent information in order to pick solutions that are appropriate, which is seldom the case. In many situations the challenge is more to know about the tools than the fact it doesn't exist. However, maintaining such transparent and neutral information is a huge effort, as CartONG and iMMAP have experienced through our NOMAD project (platform to share information about existing mobile data collection solutions, that was decommissioned in 2016 given the time-consuming process of updating the information and asking the providers).

> THE PROBLEM IS NOT SO MUCH FINDING SOLUTIONS (THEY EXIST), BUT DEPLOYING THEM IN THE CONTEXT OF OUR PROJECTS WITH LIMITED RESOURCES AND SOMETIMES LOW-LEVEL IN IM TEAMS. – HUMANITY & INCLUSION

The identification of adapted technical solutions in fact should only come second, after defining precisely the need. This often proves a challenge for CSOs, as the (technical) needs assessment phase come before the project start, when the CSO need to fund the time spent on its core funds, and try to optimize the project design/proposal time as much as possible. Most of the time, doing a proper needs assessment is not feasible in this context, and CSOs usually either ignore the data aspect completely in their project proposals, or streamline their approach by only mentioning one (or several) tool. The only way to bypass this bias in the humanitarian project design is to have solid data strategies and processes in place before (cf. chapter 9.4).

Finally, it is worth noting that while this vision is commonly shared among data professionals, whether they work in implementing organizations or in specialized structures (the workshop organized with specialists H2H organizations lead to similar conclusions), it might not be always shared by program professionals who often still focus on finding "the" perfect technical solution without identifying all the connected stakes. It would be interesting to further research this by investigating the perspective of "non-data specialists" in CSOs (but that would be a much broader research angle, that CartONG couldn't endeavour alone).

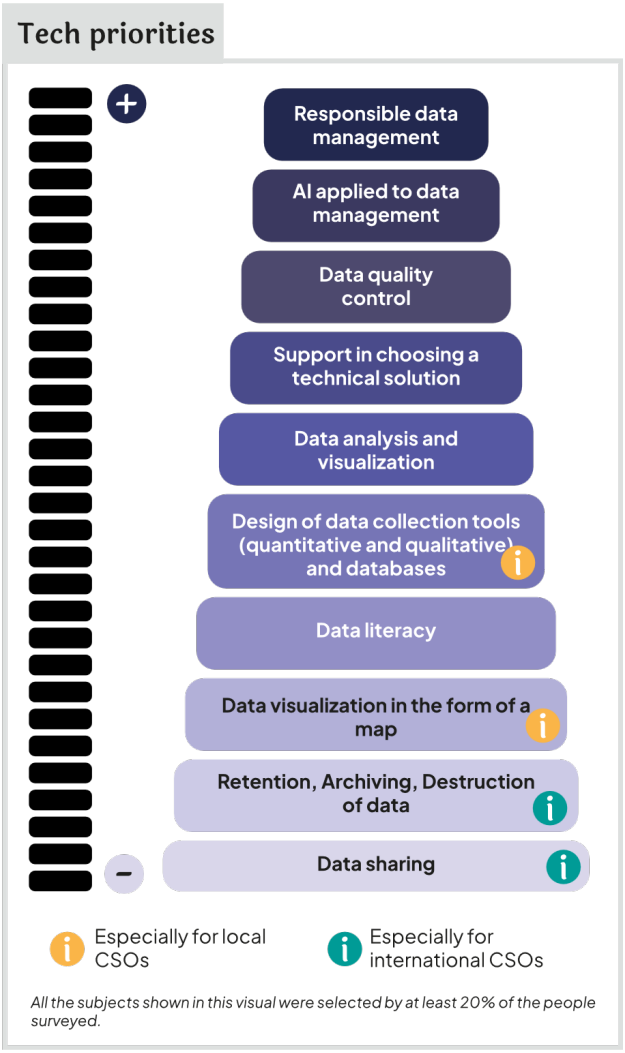
Similarly, as some tools widely used in the humanitarian sector (for instance survey-centred MDC tools) are now very cheap and easy to set up, it can become difficult to "sell" to donors or CSOs top management the need to allocate significant budget lines for other type of tools, in particular ones with more advanced features (or stronger data security).

7.2. ...AND RESOURCES...

The need for documentation and resources was highlighted a lot in our 2020 study. During the consultations that lead to this updated study, while our interlocutors still mentioned some specific expectations, the overall need seemed to be less critical. We can of course hope CartONG's intensive resource production during the period, as well as the efforts of many of our peers, have helped – although this result is biased as the respondent often come from our existing network of partners.

> IT IS IMPORTANT TO EMBED IM-RELATED CONTENT IN OTHER RESOURCES (MEAL, PROJECT MANAGEMENT…), IT IS NOT ONLY A STAND-ALONE TOPIC – EMMANUELLE LACROIX, DISASTER READY

Nonetheless, our respondents highlighted topics on which they felt support and resources could be increased. These include some "essentials" foundational to building data skills and quality as discussed in previous chapters 5 and 6 (quality control, picking the right solution, dataviz, mapping…) but also beefing up their skills on the key stakes of the sector discussed in chapter 9 (responsible data management & Retention Archiving Destruction, data literacy…) and on new topics (IA). International CSOs in particular focused on these last 2 categories, requiring resources mostly on new topics while the "essentials" seemed covered (in fact, 15% of them even mentioned not requiring any new resource!). CartONG will of course adapt our resources productions to these needs in the coming years.



We however witnessed a strong difference between international and local CSOs: local organizations reported a strong need of resources on almost all subjects, and in particular on “basics” such as design of tools or mapping. As witnessed by CartONG and other support organizations, it is harder to reach local CSOs due to their diversity and dispersion, and local support providers that could disseminate and expose the existing resources are often still lacking. The question of translation is also critical: while many resources in French have been developed

in the past years (partly due to CartONG's effort, and the French Development Agency support), the effort still remains to be started in most local languages.

For these local organizations in particular, but also more generally, the focus in the future should therefore not as much be to build new resources but to streamline the accessibility of quality and relevant ones. As a respondent put it, "overall, too much noise hinders the search for reliable and trustworthy resources". This concurs with CartONG previous researches:

> CAPACITY BUILDING IN PROGRAM DATA MANAGEMENT CANNOT BE ACHIEVED SOLELY THROUGH THE DISSEMINATION OF RESOURCES, NO MATTER HOW SPECIALISED AND RELEVANT THEY MAY BE, BUT REQUIRES A REAL VISION THAT IS REFLECTED IN AN OPERATIONAL STRATEGY WITH DEDICATED RESOURCES – AT THE VERY LEAST – IN THE MEDIUM TERM. – CARTONG[48]

Several actors also mentioned the need to contextualize these resources, to the humanitarian & development sector and different types of practitioners, to connect clearly how they can be applied on a day-to-day basis, and on "how the use of data can actually affect people's lives" (Bárbara Paes & Lesedi Bewlay, The Engine Room). Awareness of existing resources & guidance to help local organizations in particular find resources is thus especially important.

How do CSO experts find support on data management?
According to our survey respondents, data experts have multiple ways to find resources, but search engines are still the most selected approach (used by 60% of them!), before the internal documentation of their organization, external colleagues, community of practices, and specialized websites (all used by between 30 and 50% of our respondents).
Local CSO experts have a higher tendency to ask advises to their peers from other organization and specialized external structures or networks, while international CSOs more often have an internal resource person, or know where precisely to look on specialized websites.

[48] [In French] Améliorer les pratiques de formation en Gestion des Données Programmes : constat & actions pour les acteurs de la solidarité internationale, Maryline Chabanis & Marion Chranuski, CartONG, March 2022

For those using existing specialised websites, many mention CartONG's IM resource portal[49] and its Learning Corner[50], with the second big source of information being technology-related user forums or documentation (ODK, KoboToolbox, Excel, PowerBI, QGIS, SurveyCTO, CommCare, ActivityInfo, DHIS2, Bahmni…), and some mentioning United Nations websites (such as the Centre for Humanitarian Data[51]) as well as specialised websites or communities of practices (around evaluation, learning…such as the ALNAP HELP library[52], EVAL forward[53] , MERL Tech[54]). As often this figure is to relativize as our panel include many partners of CartONG, but it's still interesting to identify what are the "main" other platforms identified by our respondents.

Experts from the H2H Network, as supporters and often trainers of CSOs, had a slightly different but complementary perspective: from their point of view, the key need was to produce more practical guidance, rather than theoretic approaches, for instance on AI, data literacy, or data standardization. This study of course doesn't fill this need, but CartONG will keep this warning in mind for our future work.

On top of this need to adapt resources and their dissemination, our survey's respondents also mentioned the need for time – and this funding – for capacity building efforts, and in particular for practitioners to participate to trainings, to ingest resources, and to do tech watch. This was already identified as a key blocker in CartONG's study on the enablers of capacity building for data management[55].

[49] "Information Management Resources Portal", CartONG
[50] "Learning Corner", IM Portal, CartONG
[51] "The centre for humanitarian data", OCHA
[52] "Humanitarian Evaluation, Learning and Performance (HELP) Library", ALNAP
[53] "EVAL Forward – Evaluation for Food Security, Agriculture & Rural Development", EVAL Forward
[54] "MERL Tech", MERL Tech
[55] [In French] Améliorer les pratiques de formation en Gestion des Données Programmes : constat & actions pour les acteurs de la solidarité internationale, op. cit.

> CAPACITY BUILDING IS HARD TO FUND AS IT'S A SUPPORT FUNCTION THAT – HOWEVER MEANINGFUL – IS SEEN AS A BIT OF BLACK HOLE, HARDER TO GET A STRUCTURAL PROGRAM ON THIS – SPECIALIST FROM A DONOR

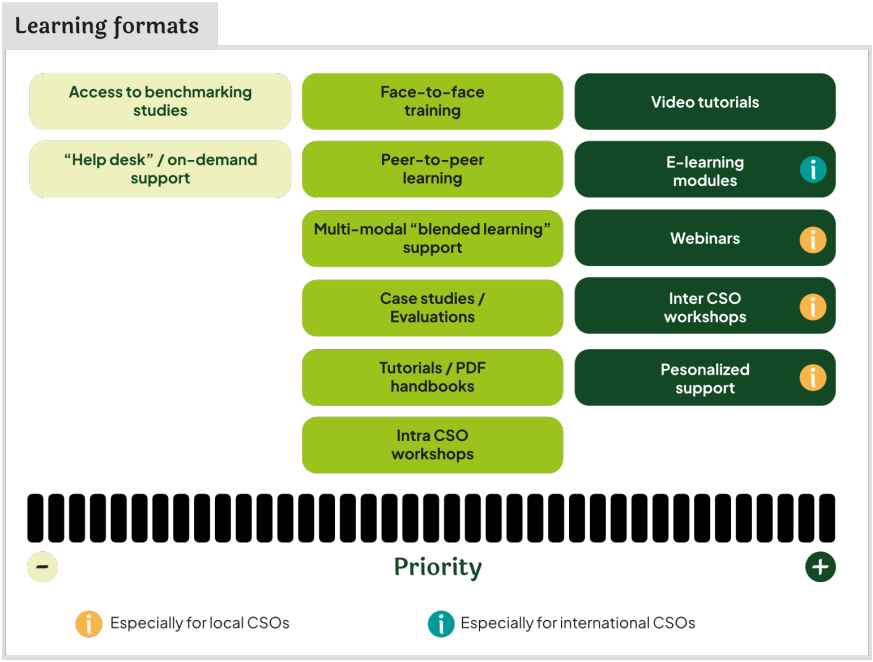
Few organizations have the possibility to invest their own funds on data management, at least at the level that they would require to face the current challenges of the sector (see also chapter 8 on how this can create a divide between the organizations that can and those who cannot). Local and smaller CSOs most often don't have the necessary financial resources, leaving a responsibility to donors and top-level organizations (big CSOs, UN agencies…) to also support a wider capacity building effort in the sector. A specialist in CSO's capacity building confirmed this specific challenge of small and medium organisations to fund a strategic approach around data that would include a proper needs assessment and identification of relevant solutions before projects start.

Of course, this question connects to wider strategy and governance aspects (cf. chapter 9.1) in terms of funds and skills distribution within the humanitarian system. As one of our respondents summarized it:

> THERE ARE ALSO QUESTIONS OF GOVERNANCE, THE INJUNCTION TO GIVE A GOOD IMAGE OF PROGRAMS, THE ABSENCE OF SPACES FOR CRITICAL DEBATE [AND…] THE OVERLOAD OF WORK ON OPERATIONAL ASPECTS, WHICH DETERS EVEN THE MOST WILLING – SURVEY RESPONDENT

7.3. ...THEN CHANNELLED THROUGH VARIED LEARNING FORMATS

Linked to the question of resources, our survey respondents highlighted that the needs in terms of training and learning are very diverse. The top formats for support & learning they mentioned are:



Most of the formats we suggested in our survey were picked by various participants (between 25% and 50% of respondents for the list above), only benchmarking of solutions and help desks being picked by less.

Generally speaking, we can observe that all CSOs are interested in formats compatible with remote learning, proving the impact of the Covid-19 pandemic on digitalisation of processes. Our respondents are also quite open to self-learning formats, which is compatible with the general uptake of skills discussed in chapter 6. International CSOs in particular put a focus on e-learning, reflecting the need to find solutions for a growing number of staffs interested in data on various locations, also in a context of high turnover in the sector. Local CSOs, on the other hand, emphasize their interest for inter-CSOs experience sharing and workshops (and webinars), which exist more in international organizations spaces indeed. As one

donor interviewed put it, ensuring individual mentoring and peer-to-peer exchanges also helps local CSOs build the confidence they need, beyond just the skills required.

That being said, offering (self) training tools and peer-to-peer learning is no substitute for proper training curriculum, professional recognition and capacity building strategies. Data management is not an exception in the professional landscape of the humanitarian sector: as found in the Bioforce institute's State of Humanitarian Professions[56], while there are competencies that distinguishes humanitarian work from non-humanitarian professions, we still face a lack of professionalisation and competencies verification. This is also true for information management, "perhaps the newest and most rapidly growing area of humanitarian work. Professionals in this area seem reluctant to identify it as a profession, yet some competency frameworks appear to have been established quickly".

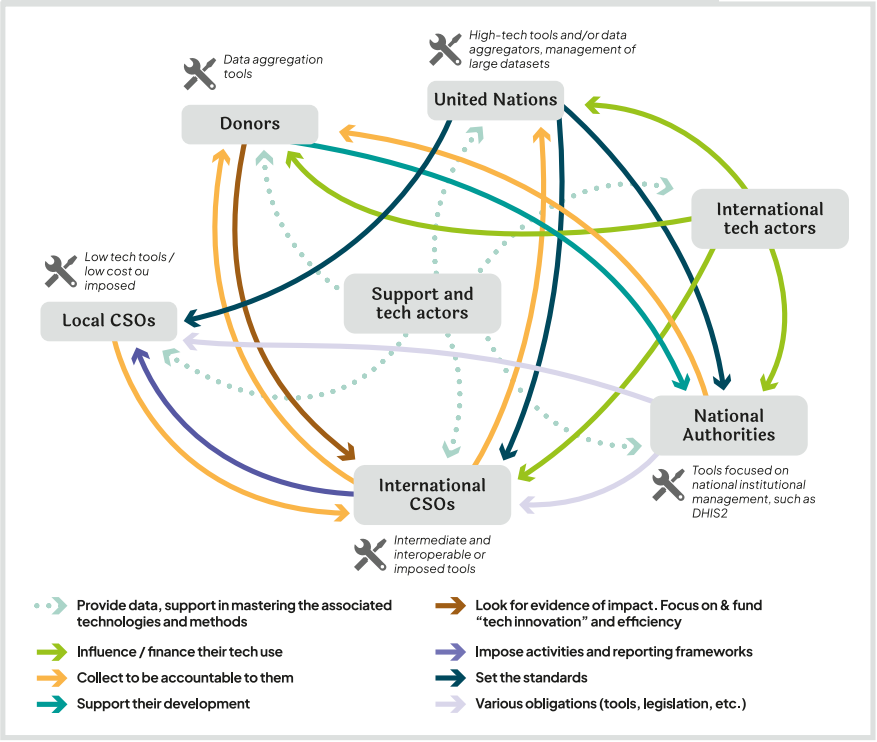
[56] The State of Humanitarian Professions, Bioforce, December 2020

8. LE RISQUE D'UN SYSTEME A 2 VITESSES POUR LES DONNEES HUMANITAIRES

As explored in chapter 5.2, the current system of humanitarian data collection potentially leaves behind large parts of the population in need of help. But are the same factors not creating the risk of a gap between the humanitarian *organizations* too – and CSOs in particular?

The humanitarian data ecosystem in which CSOs evolve is complex. We have nevertheless tried to summarize it (which some unavoidable shortcuts, in particular local communities who are the source from most data and often ignored, as explored in details in chapter 9.2), which will help understand not only the reflections in this chapter, but also in the following chapters.

Program data management at the center of a complex ecosystem



Among this landscape, there is a strong gap in capacities between international CSOs and local/national ones in the field of data, similar to the gap on other organizational dimensions – and also similar to the gap between Global South and Global North National Statistic Offices, for instance. These aspects are illustrated with various sources in our previous study "Changing the outlook: for a local approach to data"[57]. We will just add here a case study that was not sourced in our previous work and that summarizes most of the issues discussed in this chapter.

Focus: who owns data in Somalia?

A remarkable example of this inequality of the "aid data business" has been analysed by the Somalian agency Somalia Public Agenda in a policy brief[58]. Following the development of remote monitoring by international actors, a growing market of "aid data" with local providers have developed in Somalia since 2011, thus creating a "hierarchical division of labour". In this ecosystem, international organizations are the only commissioners of research, who contract international consulting firms who lead the surveying. Local organizations, researchers and consultants are mainly confined to a primary data collector role. Then "the study design, analysis and writing is then typically done by these international organizations and experts, leading to a situation in which Somalis rarely get to write and represent their own society". Most of this data being used for donor accountability, it is then not available publicly or for local research. Another side effect is the survey fatigue developed by the communities, possibly leading to data quality issues when they learn how to use this system to get more aid, on which they depend (and who could blame them?). This destroys the link between researchers and communities, basically defeating any possibility for accountability.

The "commodification" of data also is a deterrent factor for data sharing, as the international consulting firms have no incentive to share it rather than selling future markets of data collection for the same information. This is also linked to the general lack of use of data for decision-making, with data used much more for donor reporting and thus not shared between operators (cf. also Chapter 9.3).

[57] Changing the outlook: for a local approach to data, op. cit.
[58] "Who owns data in Somalia? Ending the country's privatised knowledge economy", op. cit.

The case study concludes on the high necessity for a state regulation, both on data protection but also on forcing data standardization, sharing (including with open data), and involvement of local actors (and in particular researchers/data specialists) in analysis and definition of research object: all these recommendations will appear in different sections of our report!

We haven't identified a lot of research on this topic so far, however both CartONG's field experiences and several testimonies collected during our interviews for this study highlight the risk that this "hierarchical division of labour" could become more pronounced in the future.

> ONCE A SURVEY HAS BEEN CARRIED OUT IN THE FIELD, IT IS SOMETIMES NECESSARY TO WAIT FOR AN INTERNATIONAL CONSULTANT TO COME AND PRODUCE THE ANALYSIS – A LOCAL CSO

CartONG – and several specialists interviewed or read during the preparation of this study – therefore identify a strong risk for the humanitarian system to break into "two tiers"[59]. This portrayal is of course simplistic, with many organizations (in particular the "medium-sized" international CSOs) falling rather in the first or second category depending on the topic. But we thought it would be interesting to summarize these 2 categories in the following table:

Typical profile	UN organizations, major international CSOs, donors, National Statistics Offices in Global North	Smaller international CSOs, most of national/local CSOs, most National Statistics Offices in Global South
Resources	Enough funding relatively speaking (access to indirect costs, etc.)	Scarce resources with underfunding from institutional funding or lack of core revenue

[59] The vocabulary is inspired by articles who classify CSOs in "tiers" depending on their size, wealth and influence; for instance our "major CSO" first tier would correspond to the "Tier 1 NGO" in this article, while Tier 2 and 3 would rather fall in our second category, or in-between: "The big list of humanitarian NGOs", The humanitarian insider, February 2023

Skills	Up-to-date skills and relative capacity to compete with the private sector for talents	Lack of skills, loosing top staff to bigger organizations
Innovation	Capacity to increase productivity with new technologies, in particular AI	Suffer technical evolutions without resources to take control of them
Ecosystem	Enough weight to negotiate with global providers / Big Tech and co-build or challenge their solutions	Dependency on global solution providers without negotiation leverage
Type of data expected	More global but standardized, usable for high-level decision-making (fund allocation)	Higher granularity, usable for day-to-day program decisions
Standards and ownership of data	Setting up high standards to maximize comparability, and store in global repositories	Trying to push localized standards and own data at local level
Place in the data chain	Defining data needs and methodologies, analysing it and using it for decision-making	Collecting data in the field with limited opportunities to analyse it or build the research agenda
Main objective	Accountability to donors / tax-payers	Impact on communities

This summary in 2 categories is of course stereotypical, it will however be useful to keep it in mind while we discuss several structural challenges of the sector in the next chapters.

The link between CSOs and the data private sector is an important component dimension of this, since innovation and technologies generally are issued from them, as reviewed by Access Now in a comprehensive recent research[60].

> IN THE HUMANITARIAN TECH FIELD, A HANDFUL OF BIG TECH COMPANIES COMPETE FOR THE RELATIVELY MEAGRE TECH FUNDING AVAILABLE WITHIN A SMALL SUBGROUP OF 'BIG NGOS' […]. MEANWHILE, ACTUAL CAPACITY TO SUCCESSFULLY ADOPT AND ADAPT TECH WITHIN INTERNATIONAL AND LOCAL NGOS REMAINS LIMITED, AS THESE ARE STILL GENERALLY LOW-TECH ENVIRONMENTS. – ACCESS NOW

Access Now illustrates well the divide between the two levels, with smaller and more local actors struggling to both cope with the same heavy accountability requirements as larger ones, while having at the same time to face the technology challenge. This also creates risks of dependencies to solutions offered by Big Tech companies even for larger organizations that have access to them, as detailed in chapter 9.12. In fact, there is a permanent risk that even international CSOs that are big in their country but still relatively small at the global level (such as most French CSOs for instance) get "second-tiered" when having to negotiate with Big Tech companies, so this division risk should be everyone's concern.

Development Initiatives, a specialist organization working on these issues in particular through the statistical angle, has warned on a recent review of the UN's Global Digital Compact on the illusion that technology could solve the digital divide[61].

> THE IDEA THAT TECHNOLOGICAL INNOVATION AND BIG DATA ARE THE SOLUTION TO LOW-INCOME COUNTRIES' DATA PROBLEMS IS CURRENTLY GAINING CURRENCY IN MANY DEVELOPMENT CIRCLES. AS WE HAVE ARGUED BEFORE, THIS IS A DANGEROUS CONCEPT, AND THIS

[60] "Mapping humanitarian tech", Giulio Coppi, AccessNow, February 2024
[61] "Digital Compacts: Global ideals, regional realities", Stephen Chacha, Bill Anderson, Development Initiatives, September 2024

> LEAPFROGGING IS UNREALISTIC. – DEVELOPMENT INITIATIVES

A caricatural example of this risk of two-tier system, that goes beyond the humanitarian sector itself, is WorldCoin: this start-up (created by Sam Altman, founder of OpenAI) made the promise of universal access to their new cryptocurrency through the use of biometrics[62]. Building on the financial and technical resources of Silicon Valley, WorldCoin encourages the registration of new users by distributing small amounts of cash when registering one's biometrics. However, it turns out that a large share of their users are insolvent, fragile populations that see this small revenue as an opportunity, with intermediaries that push them to register while keeping a margin. So WorldCoin basically buys sensitive, private biometrics data from people that have no understanding of their commitment. Currently a quarter of their total 4,5 million users are from Argentina, building on the economic crisis and lack of cash of the local economy… while the platform is not allowed in the USA and has been recently banned from Spain.

To ensure the benefits of digitization and data for Global South countries, Development Initiatives highlights the need for solid infrastructures, basic data literacy spread out, and working field data systems to allow the possibility for innovation to strive. They quote the African Digital Compact, that lists as necessary conditions "bolstering digital infrastructure, enhancing cybersecurity measures, fostering digital literacy, implementing robust data protection laws, nurturing local technology ecosystems, modernizing regulatory frameworks, and deeply engaging community stakeholders".

Bridging the digital and data divide to make sure there is one unified data ecosystem in a country as a key evolution, needed in the sector. Drawing from multiple field experiences, they recommend concentrating investments on field actors and foundational datasets: local administrations (civil registry), primary schools and health centres, etc. The goal is to build a backbone for data management in a country, which can then be shared (but not prescribed by) to global organizations[63]. In fact, interviewees from donors testified that similar

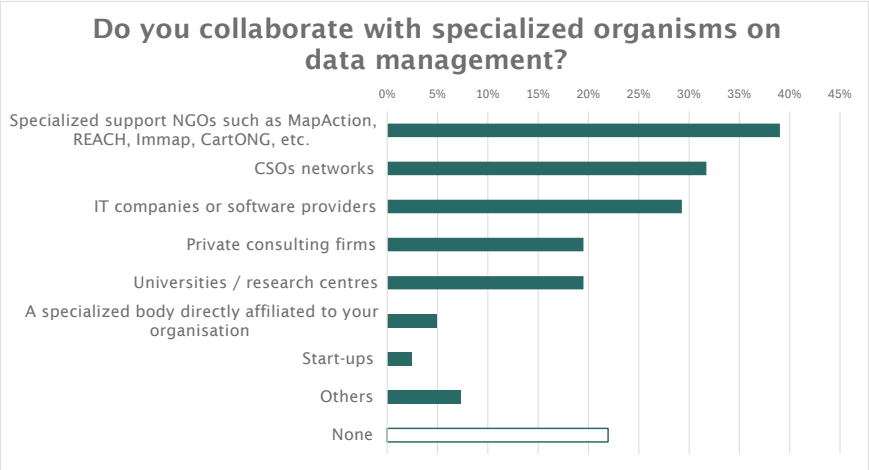
[62] [In French] "De l'argent contre des données biométriques : la start-up américaine qui profite de la misère", Louise André-Williams, Médiapart, March 2024
[63] "The data side of leaving no one behind", Bernard Sabiti, Bill Anderson & Sam Wozniak, Development Initiatives, September 2021

challenges are faced by governmental institutions and in particular National Statistics Offices, which pleads for a country-wide ecosystem vision (cf. next chapter 9.1). There is a temptation for donors to only fall back on supporting local actors (which is of course a necessary first step) and giving up on more structural transformations of the systems, which are necessary harder and even sometimes out of reach for them.

But success stories and case studies show that the alliance of different type of actors can bridge this gap. The DHIS2 platform, a success story deployed in more than 60 countries, is a good example of such systems. READY (Global Readiness for major disease outbreak response) illustrated the case study of the Democratic Republic of Congo, where many international partners have supported various Congolese public institutions, including the Integrated Analysis Cell (CAI) within the Ministry of Health since the Ebola epidemic of 2018-2020[64]. The CAI plays different roles including analysing, sharing results, and steering actions. All this while leaving control of the data to local field actors in health facilities, who can use their own data to understand health dynamics and identify the causes of epidemics for prevention purposes.

Another part of the solution to “bridge this gap” could be to work with specialized, support organizations. For instance, most of the H2H Network members now have a strong positioning towards supporting local actors, creating equity along the data production chain, and promoting responsible and sharing practices. 78% of the CSOs who answered our survey in fact rely to some extent on one or several specialized partners on data: the first two categories being support NGOs such as MapAction, REACH, iMMAP, or CartONG, and networks of CSOs, but with a variety of other types of partners (cf. graph). These “middlemen” of humanitarian data management can be key agents of change in the transformation of the system.

[64] “Why the delay? Perspectives of national and local actors on progress toward locally led outbreak readiness and response”, READY Initiative/Anthrologica, October 2023


Do you collaborate with specialized organisms on data management?
0%
5%
10%
15%
20%
25%
30%
35%
40%
45%
Specialized support NGOs such as MapAction, REACH, Immap, CartONG, etc.
CSOs networks
IT companies or software providers
Private consulting firms
Universities / research centres
A specialized body directly affiliated to your organisation
Start-ups
Others
None

9. A CHANGING FIELD: CURRENT AND FUTURE STAKES

CHAPTER METHODOLOGY AND INTRODUCTION

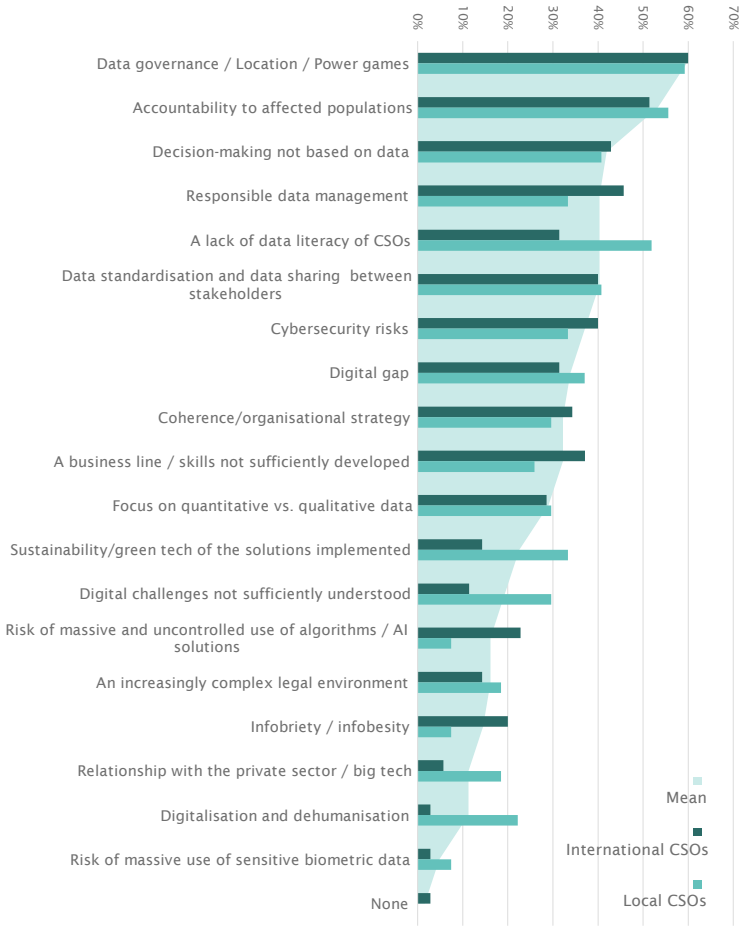
Even though the sample from our survey has its limitations, given the limited number of global surveys for humanitarian data specialists to express what were top challenges of the day for the sector, this chapter's order has been built on our survey's responses, completed through various interviews and sources.

To be consistent with our previous work, we have again retained 12 stakes with 8 priority and 4 secondary ones (in 2020 we retained respectively 6 and 6). Among these 12 stakes, 8 were already there in 2020 – we are thus developing on how the associated debates evolved in 4 years – and 4 are new – either because they weren't seen as priority, or not that identified back then. We have also ordered them to have logical sequencing and avoid repetitions. The ranking is thus not necessarily exactly the same than our survey respondents'.

Stakes in program data management

Four stakes that were well identified by our respondents are not represented as a chapter here, because we developed them in other chapters of the report (structuring a professional field in chapter 6; focusing on quantitative data rather than qualitative data in chapter 5.2; digital divide in chapter 8; and info-sobriety & infobesity with stake #11).

Finally, 3 topics didn't receive much vote (digitization and dehumanization; relations with the private sector/Big Tech; risk associated with massive use of big data and in particular biometrics) even though all these specific questions are touched upon in various chapters of this study.



When going into a bit more detail, we can also notice that while the key challenges are equally identified by local and international CSOs, some specificities appear: international CSOs are more concerned by more "macro" and new challenges such as infobesity and AI in particular. Local CSOs are proportionally more concerned with challenges that concern directly the people behind the data and control over systems: dehumanization of processes, relations with Big Tech, biometrics, sustainability of systems, and digital stakes in general. But we could also assume that some of these more strategic (and less practical) stakes seem less of a priority for international CSOs because they feel more in control of them than local CSOs – we'll illustrate this in the next chapters.

9.1. DATA GOVERNANCE & LOCALIZATION

As detailed in our previous study "Changing the outlook: for a local approach to data"[65], while localization has been a top priority of the humanitarian agenda since 2016, most actors and in particular local ones consider it is lagging. In fact, since 2022 the concept is questioned with a new approach pleading for a "decolonization of aid": in reality, these two angles work in parallel, with "a technical discussion about how to make aid better, and a moral conversation on how to address the wider geopolitical power dynamics that led countries to be in need of aid in the first place"[66].

These debates have important implications for the way data is managed in the humanitarian sector, in terms of governance, allocation of resources, but also methodologies and choices of tools – data being seen alternatively as a technological tool to solve localization challenges, or on the contrary as a technology perpetuating colonial oppressions. This question of data governance, and behind it, data equity, has – as it often is in the humanitarian sector in general – indeed been identified by our survey respondents, both local and international CSOs, as the top issue our sector needs to tackle.

The question of empowering local actors on program data management is of course deeply intricated with almost all the other stakes that will be covered in this chapter. We have already discussed how the lack of inclusivity of data collection processes connects to severe biases in data quality (cf. chapter 5), linked to the

[65] "Changing the outlook: for a local approach to data", op. cit.
[66] "Policymakers and racial justice activists came together to discuss decolonising aid. Here's what happened…", Heba Aly, The New Humanitarian, August 2022

very professional structuring of the field (chapter 6), and on the fast technology advances create the risk of increasing the gap between actors (chapter 8).

In fact, this change in power relationships between actors cannot but be connected to the question of data governance (understood as data governance of the sector, not within a specific organization – for the latter please refer to the comprehensive documentation developed by NetHope[67]). While being aware of not piling up concepts that can be hard to grasp for small and local organizations, it is critical to understand that these challenges are only the consequence of the very organization of international solidarity. The different deterrents to equitable partnerships identified for overall localization efforts of course apply in the data branch: the current project management models builds on interests and fears of international organisations and their staff, organisational cultures, structural barriers, inequalities of power, and lack of trust.

Another important aspect to keep in mind is the diversity of situations. The Ukrainian crisis is often mentioned by professionals as the counterexample where localization commitments were achieved. In a specific context of high availability of resources and technical skills, local CSOs have managed not only to take a leadership role in the humanitarian response, but also to integrate the reporting (and associated data management) requirements by the international system. A case study on local CSOs in Poland have shown the importance of peer-to-peer capacity sharing (rather than formalized, and sometimes repetitive and unfit) formal trainings in organizational development for which international CSOs are not professionals[68].

However, these differences can also have negative ripple effects. Indeed, as a specialized organization testified, in many “forgotten crisis” (e.g. Central African Republic, Chad, Burkina Faso, etc.) the lack of available resources (and higher practical complexity to collect data, e.g. harder to get internet or mobile devices) hamper the capacity to collect data, which in turns lead to less evidence, less funds allocated and less available resources, creating a vicious circle.

[67] “Data Governance Toolkit: A guide to implementing data governance in nonprofits”, Micheline St. Clair (Plan International), Abhijit Balakrishnan (SOS Children’s Villages); Morgane Bradley (NetHope), NetHope, February 2022

[68] “Grand Bargain Localization Commitments (Poland Case Study)”, Dominika Michalak, Véronique de Geoffroy, Rana Gabi, Elie Keldani, Karina Melnytska, NGO Forum Razem & Groupe URD, June 2024

From our previous work, we have identified a series of challenges and associated recommendations:

- letting local actors define the type of support they need, and help develop their skills (including training but also equipment), in particular structuring local capacities on base datasets (cf. chapter 8);
- tackling the biases in data representation (cf.5.2) and the project management system that leads to it, in particular through better recognition of the role of local authorities and National Statistics Offices, and exploring new project management methods leaving more space for co-construction with local actors;
- working on the inclusiveness of the data chain, through participation of communities in how they are represented in data, data literacy democratization, data sharing and open/responsive systems (this list is coming from the Data Values project manifesto[69] but most charters on the topic follow similar principles);
- working on data ecosystems, both at the local level (cf. chapter 8), but also globally through distributed leadership networks (such as the OpenStreetMap movement or Flying Labs network);
- invest in the humanitarian experts of tomorrow to multiply local data initiatives and build skills that will help correct the biases observed.

If we focus more specifically on the perspective of CSOs, some challenges are shared by all of them, in particular the perspective of a 2-tier data humanitarian system (cf. chapter 8) where most CSOs except a handful of major ones share the same risk of being relegated. However, the capacity to act of CSOs obviously depends a lot on their position. For international actors, when exchanging with francophone CSOs[70], they identified mostly challenges linked to appropriation of the topic by local actors: getting local authorities involved, enforcing local data regulations (on top of international ones such as GDPR), getting local partner involved throughout the projects and in-between them, etc.

Some development oriented international CSOs are very involved at a partenarial level into working on data both useful and usable for local partners (example of Acting for life). International CSOs often struggle with the diversity of the skills of

[69] "The #DataValues Manifesto: Demanding a fair data future", Data Values project, 2021
[70] Program data management seminar, June 2024

their local partners in terms of data skills, even the willing ones having difficulty involving their partners for instance on designing data collection exercises, or on organizing restitutions of the data collection results. The lessons learned from their efforts concur with the recommendations listed above: reinforcing skills and harmonization of practices; defining strategies and sharing agreements between stakeholders, including national authorities, and identify focal points capable of defining the legal framework applicable; put local partners as co-manager of data from the start.

The local CSOs who responded to our survey are all expressing their interest on growing their capacities, as well as the need to give them more funding and support to do so (quite consistently with small international CSOs in fact, confirming the risk of "two-tiers" system already discussed). Some mention the first positive effects of localization programmes, a Congolese CSO for instance mentions capacity building efforts by OCHA to train local organizations on mobile data collection as part of a crisis response. But in general, from their perspective, much remains to be done.

In general, advancing on localization will mean significant changes in the governance of the sector, and on the positioning of international organizations, whatever their status. While national governments also take a more strategic role in the organization of the data landscape of their country (including with emerging national data clouds, new data protection regulations, etc.), international CSOs will have to navigate between preexisting constraints from their international supervisors (donors, their HQ countries' governments, UN), new legitimate constraints by national authorities, and of course a reinforced dialogue with their local partners.

Finally, while we focus on this chapter only on the governance of the humanitarian sector itself, it is essential to keep in mind that the technological ecosystem in which it evolves has critical impact on it: this will be detailed in the secondary stakes 11 (on sustainable technologies) and 12 (on mastering digital stakes).

9.2. ACCOUNTABILITY TO AFFECTED POPULATIONS (AAP)

In parallel and directly connected to the question of data governance (which is looking more at the organizational or system-level), the involvement and accountability to the communities is a critical component of the reform of the humanitarian data system – it has been identified as the second most important by more than 50% of our survey's respondents (both from international and local CSOs), next to the connected topic of localization we just covered.

As defined in the latest Core Humanitarian Standards[71], accountability is the "process of using power responsibly, and taking account of and being held accountable by different stakeholders, primarily those who are affected by the exercise of such power. It means putting people and communities at the centre of decisions on issues that affect them". This definition applies also to data management: having control over one's data is a way of exercising power. In the sphere of data, accountability thus implies responsible and transparent data management, putting people and their expectations at the centre of organisations' concerns.

Without going into detail concerning the reasons (documented widely elsewhere), one can only witness that the "participatory revolution" that was one of the 2016 "Grand Bargain" commitment is 8 years later still far from the mark, with more "participation-washing" than real involvement from local populations (and their representatives) throughout the humanitarian project cycle.

In the data sector, this can be illustrated by the fact – that has been valid for a decade at least – that data is collected and used disproportionately for accountability to donors rather than towards the populations CSOs are trying to serve (cf. chapter 5.1).

The Somali Public Agenda case study we already highlighted provides a critical illustration of this:

> A LOCAL RESEARCHER FAMILIAR WITH THE AID DATA BUSINESS CAME TO THE CONCLUSION THAT "LOCAL COMMUNITIES WERE THE LOSERS OF AID INFORMATION

[71] Core humanitarian standard on quality and accountability, CHS, 2024

> GATHERING BECAUSE THE SO-CALLED RESEARCH ABOUT THEM IS NOT ACTUALLY ABOUT THEM. HE AND OTHERS HIGHLIGHTED THAT THE ENTIRE RESEARCH AGENDA – FROM THE RESEARCH DESIGN TO THE QUESTIONNAIRE AND ITS IMPLEMENTATION – WERE DECIDED ON AND DRIVEN BY EXTERNAL ACTORS. A STAFF OF AN INTERNATIONAL NGO ECHOED THIS SENTIMENT WHEN HE HIGHLIGHTED: "THE LOCAL COMMUNITY, WHICH HAS BEEN THE SOURCE OF INFORMATION AND THE FOCUS OF RESEARCH DOES NOT USE THE RESEARCH REPORTS AND FINDINGS. THESE REPORTS ARE NOT EVEN DESIGNED FOR THE LOCAL COMMUNITY TO USE IT, THE LOCAL COMMUNITY 'WAA LAGU AWR KACSADA' [THEY ARE JUST USED TO ACHIEVE THE PRE-DETERMINED PURPOSE]." – SOMALI PUBLIC AGENDA

This case is a clear example of local communities' lack of ownership concerning aid information, here in Somalia and Somaliland. It is particularly obvious in the lack of dissemination of research findings to local communities. The results of surveys and studies are usually presented in closed door donor meetings in Nairobi, which are predominantly attended by expatriates. This illustrates how outside actors – international organizations, private companies, and research outfits from the global North – dominate not only aid data production, but also aid data ownership.

There can be hurdles for CSOs striving for better accountability to affected populations beyond these system constraints. First off, there is a strong need to adapt to affected populations' data literacy knowledge to ensure that they are in a position to get involved, adapting and translating the project cycle methods of CSOs into day-to-day meaningful issues. Beyond that, there can be biases that get in the way. For example, the question of who are the "local populations", or at least who represent them as part of CSO programming and its accountability mechanism, can be questionable: as mentioned by Development Initiative, local CSOs, although they usually have a very strong contextual understanding, do not necessarily represent all the needs of "affected populations" – they usually have their own agenda, and are part of a wider ecosystem that also needs to be considered, with local authorities, village representatives, other civil society actors… which adds layers of complexity.

As discussed during a seminar of data specialists among francophone CSOs[72], accountability from an international CSO perspective often corresponds in practice only to making available feedback & complaint mechanisms and the frequent (but not constant) sharing of data results after a data collection occurrence. CSOs however testify that they are getting better in at least making the data presented to affected populations more understandable. Which doesn't mean there is still work to do to ensure the associated explanations & sharing are also appropriate. But obviously this is just a drop in the ocean of what it should be in terms of level of involvement for proper accountability: there should be strong involvement and leadership of communities' representatives in the definition of program aims and therefore expected results, to ensure its relevance.

It is worth noting the UN/public system is facing the same challenge: while there have been strong progresses in high-level principles such as open data (for instance the UN Statistical Commission adopted "open data by default" in 2022), it is still difficult to prove a link between data sharing and citizen accountability[73]. Accountability is possible only if shifting from data availability to data use by citizens, i.e. if understandable and relevant outputs are available, whether it is through data visualization or publication such as media articles, and of course through translation in local languages (cf. chapter 5.2).

Accountability to affected populations can be seen under different data lenses that are all connected to each other: more involvement of local populations in data processes, more data quality, more data security and more transparency related to the data exercises that inform programming.

First off, there is the question of involvement – although we can see that there have been small improvements, international and local CSOs need to be a lot more proactive in engaging local populations or their representatives in the definition of project activities (and therefore their translation into data needs), the associated data analysis and sharing, as well as project learning. The conditions under which CSOs prepare their projects (with generally very short application deadlines from donors, and no funding for preparatory work) rarely allow this work to be funded, similarly to the identification of relevant technical solutions (see Chapter 7.1). Of course, this will never guarantee participation (CSOs cannot force people to do

[72] Program data management seminar, June 2024
[73] "Testing the Assumptions of the Data Revolution", op. cit.

so!) but there are multiple tools and approaches that create the necessary conditions to foster participation. For instance, new methods of project management such as the change-management approach developed by F3E[74] can be mobilized to do so. In fact, some donors are beginning to finance these preparatory phases for development projects, an approach that can only be encouraged for the whole aid industry.

Innovative data approaches and technology can offer many possibilities on this front (participatory data collection methods to evaluate needs uphill of a project, better involvement of populations in data exercises / access to their data, a virtuous data sharing cycle that help with project improvements, etc.). Feedback and complaints mechanisms are also more and more present thankfully (at least for international CSOs, as seen in chapter 6.2), but can be seen as curative rather than preventive measures for improvement and quality of action and should therefore never be the only method used for participation. These different types and level of possible digital participation are also described by Groupe URD in their study on “participation and digital technology”[75].

The next step is to ensure that the data used for decision-making is relevant actioned on. It means building rigorous analysis plans that leaves space for qualitative data (as discussed in chapter 5.2) rather than going on quantitative-only approaches sometimes presented as more scientific, but that are often more limited in terms of nuanced and contextualised analysis. Beyond that, in the same way that one of the digital principles is to “design with people”[76], which means “to invite those who will use or be affected by a given technology policy, solution, or system to lead or otherwise meaningfully participate in the design of those initiatives”, it is essential to increase involvement of affected populations at all stages of the program data cycle, from the definition of needs to the sharing of results and learning. There has been a small shift of paradigm on this over the past couple of years that is promising, as is shown for instance by the growing frequency of the topic during francophone CSOs exchange days. Although a lot remains to be done, in particular in utilizing this data for decision-making and

[74] "PRISME Toolkit", F3E, November 2022
[75] “Critical analysis and key messages ’Participation & digital technologies’”, Marie Faou, Groupe URD, 2023
[76] “Design with people”, Principles for Digital Development

connecting these different data sources: for instance data from complaint mechanisms is rarely connected with monitoring & evaluation systems.

Thirdly, the question of securing the data – it remains of course a key "accountability" focus to secure the personal and sensitive data that CSOs are entrusted with to ensure it is used in accordance with their needs and does not fall into the wrong hands. This also encompasses data minimisation, i.e. collecting only the data that is really needed, to limit the risk of data breach as much as possible, while also respecting – in a human-rights based approach – the right to privacy (cf. also chapter 8.6).

And finally, the question of transparency – this is related to communicating on data-related activities as much as necessary to ensure adequate understanding by those wishing to know more, and make participation possible when the context is conducive. This means uphill of data collections planning to inform affected populations collectively of data collection exercises, as well as the more individual version of this during specific interviews/surveys/focus group discussions, before asking for respondent consent. It also means the sharing of data exercises results – in an understandable form – for feedback from and discussion with communities, as well as keeping access possibilities open to respondents concerning the data that they shared.

Focus: accountability vs. misinformation & disinformation

An emerging new threat that further complexifies accountability to populations is disinformation or "fake news". Its objective is "to alter your perception of reality in the long run. Disinformation campaigns are usually launched by foreign states, terrorists and sometimes even domestic actors, with the aim of creating a climate of mistrust. The ultimate goal for these actors is to destabilise regimes and institutions for political purposes, which usually ends in conflict and tragedies" (definition by the CyberPeace Institute)[77]. Disinformation is not limited to governments, and the humanitarian sector is increasingly confronted to it. The CyberPeace Institute takes the example of disinformation operations lead by actors (including states) during the Covid-19 pandemic, manipulating and leaking authentic

[77] "Disinformation and misinformation", Responsible data management Toolbox, CartONG, June 2022

information to undermine people's trust in healthcare organizations[78]. The raise of AI has of course reinforced the risks and amplified possibilities, through learning and reproduction, to invent information. Confronting disinformation is not an easy task for CSOs, as it requires both to monitor information channels to identify and counter it – which is mostly the responsibility of communication teams – but also a link towards accountability and thus monitoring & evaluation and data management, in particular to counter fake arguments and provide evidence about the organization's work. While this work adds a new burden to already busy teams, it is critical to maintain trust with affected populations to counter disinformation, otherwise all accountability efforts might be wiped out. For more information on how disinformation works and how to counter it, you can check Internews' toolkit[79], or the recommendations from GSMA[80].

However – like for many aspects that are connected to the "localisation" shift in paradigm – CSOs often feel that the 'accountability to donor' system they are in does not leave much space to make involvement of affected populations real. These approaches are indeed often quite resource consuming. Unless they have core funding to take initiative uphill of proposal submissions, they struggle to include budget lines for this, which is seen as "support" lines or even indirect costs, this diminishing programs budgets.

> THERE IS NO FINANCIAL INCENTIVE FOR HUMANITARIAN ORGANIZATIONS TO CHANGE THEIR PRACTICES AS LONG AS THEY ALIGN WITH THE DONORS' WISHES – NIKLAS RIEGER, DEVELOPMENT INITIATIVES

This imperative is nonetheless part of the virtuous cycle to ensure that programs are as useful and relevant as possible. And therefore, as much as possible to mobilise and make their case to donors on such topics to ensure that dedicated funding prior or during projects can be obtained to fund such approaches.

[78] "Playing with lives: Cyberattacks on healthcare are attacks on people", CyberPeace Institute, March 2021
[79] "Managing Misinformation in a Humanitarian Context", Viviane Lucia Fluck, Internews, July 2019
[80] "7 GSMA Recommendations to Reduce Misinformation, Disinformation, and Hate Speech", Wayan Vota, ICTWorks, May 2023

9.3. DATA-DRIVEN DECISION MAKING

As we have already seen, many actors criticize the fact that data is mostly produced for accountability / compliance to donors reasons, rather than to be used to improve programs. A good example is the Somali Public Agenda study already highlighted, who found out that for the vast majority of the data professionals they interviewed, data was not used for decision making outside of donor accountability: "informants described monitoring reports as a formality driven by globally adopted standards and best practices in the aid industry such as the International Aid Transparency Initiative (IATI). These standards are meant to increase aid agencies' accountability towards taxpayers as well as beneficiaries. But in reality, they are often a formality and, according to the owner of a Somali consultancy firm, few agencies substantially engage with the data and reports from the field, which they produced or commissioned."[81].

As already cited, research on the use of data for SDG measuring had concluded a low-level impact of data on decision making at political level[82]. This comes from intricate causes, including lack of access to timely and quality data (often still siloed, except for satellite data), insufficient investment on data collection, quality control and capacity building to use data with growing inequalities (including after the COVID-19 pandemic even though it raised public interest on data), etc. "Global evidence is still largely missing any rigorous analysis of the value generated throughout the data value chain, spanning from production to applications supporting decision-making nodes across government and communities."

Other data specialists confirmed it: in a context of resources not meeting the humanitarian needs and injunction to "do more with less", everyone is concerned with the question of prioritization, and data features a heavy role in these discussions. Two scenarios are open for the future of the sector:

- In one scenario, the current continuous push for more data collection, with no proper methods and strategic prioritization, will lead to a purely quantitative, massive, "tick-the-box" approach. The objective of humanitarian organizations will become to maximize the resources allocated to their work by collecting the data they feel the decision-makers in this system are expecting. If pushed to the extreme, this system will lead humanitarian

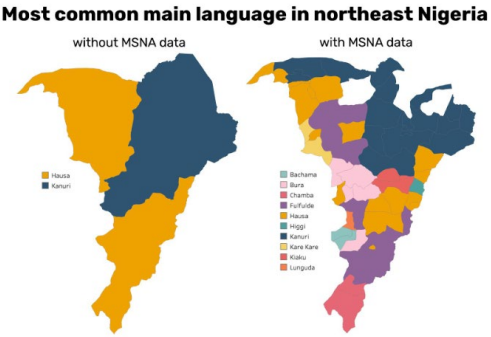
[81] "Who owns data in Somalia? Ending the country's privatised knowledge economy", op. cit.
[82] "Testing the Assumptions of the Data Revolution", op. cit.

organizations to revert to simplistic, minimal approaches, getting affected populations what they need to survive without questioning the causes of crises. Decisions will be made at global level or in capital, with decision taken remotely based on data collected.

- Alternatively, a more localized and participatory approach towards data collection, assessments and ultimately decision-making is also possible. The decision then emerges from a discussion informed by a mix of qualitative and quantitative data, rather than massive quantitative data. This scenario would require a serious shift from our current trend, raising -again- questions on the power dynamics of the sector.

All the data professionals interviewed as part of our study agree on the fact that proper data-driven decision making would rather look to this second scenario.

Several actors pleaded for a win-win scenario where data standards and datasets are built and aggregated at global level to inform high-level decision-making (such as fund allocation) while leaving some margin to field actors to collect the data they need for operations, in a way that is actually listening to the voice of the communities they ask (cf. chapter 5.1).



A good example of global tools that can be amended are Multi-Sector Needs Assessments (MSNAs). Since 2021, Clear Global/Translators Without Borders has worked with partners to integrate the language variable in MSNAs in several countries[83]. This helps understand a major gap in data quality (cf. chapter 5.2) and it's a first step for more adapted assessments via correction measures, for instance in staff hiring, training, data analysis, communication material to be translated,

[83] "How MSNA language data can improve communication with crisis-affected people", Mia Marzotto & Ellie Kemp (Translators Without Border), IM Portal Blog, October 2021

design of community feedback mechanism, and eventually integrating language into strategic planning[84].

Another interesting example is the way Doctors Without Borders/Médecins Sans Frontières (MSF) functions, one of the few (if not the only) humanitarian organization with no external compliance regulation to follow except its self-defined accountability rules. While MSF has limited data standard internally (not having IM officers positions) and leaves a lot of freedom to missions and staffs to define their data models and products, except for base data. When standardization happens (for instance for security data), standards are built incrementally based on existing data, with specialized staff integrating preexisting data and replicating standards across operations. This process is in fact quite similar to the way standards are incrementally built in HDX (cf. chapter 9.8).

Finally, the question of data-driven decision is intimately linked to the question of data literacy that we'll explore later (cf. chapter 9.5), as basic data literacy is necessary for proper interpretation of data, to avoid the biases we discussed (chapter 5.1). As a data specialist put it, "data starts the dialogue, and the dialogue informs the decision making and who should be involved".

A possible success story of data-driven decision-making: anticipatory action

Anticipatory action refers to actions taken to reduce the impacts of a specific, imminent, forecasted hazard before it occurs, or, before its most acute impacts are felt. The actions are carried out in anticipation of a hazard's predicted impacts and based on a forecast or early warning of when, where and how the event will unfold – with different approaches including more or less pre-planned trigger values and actions[85].

Anticipatory action relies heavily on data as it is based on trigger mechanisms, either via observational data (e.g. water levels) or predictive elements (e.g. flood forecasts). Designing a trigger requires access to 3 types of data: i) current and historical data about the hazard; ii) data on the historical and expected impact of the hazard and shock and iii) hazard forecast data. Availability and usability of the data remains the main barrier to

[84] "20+ language tips for effective humanitarian data collection", CartONG & Translators Without Borders, IM Portal, September 2020

[85] "What is anticipatory action?", Anticipation Hub

implementing anticipatory actions in many humanitarian contexts.
Despite this challenge, various examples have already demonstrated the possible impact of anticipatory action when data is available, for instance the Red Cross Red Crescent forecast-based financing[86] (active or being developed in around 30 countries), the START Network START Fund Anticipation[87] (which has been activated on flooding, heatwaves, cold waves, conflict and displacement, disease outbreaks, drought, volcanic activity and tropical cyclones), via bilateral facilities such as the Anticipatory Humanitarian Action Facility (WAHAFA) supported by the German Federal Foreign Office through Welthungerhilfe[88], or at governmental level (for instance in Philippines, Indonesia or India).
The anticipatory approach is gaining strong momentum in the past years. Monitoring & evaluation of anticipatory action is still limited, with need of more investment in it, common analytical frameworks and improving the models (and in particular their transparency)[89], even though it is growing[90]. It tends to appear that anticipatory is most efficient when tackling time-bound hazards rather than long-term multifactorial crises[91].
Anticipatory actions is also interesting as it brings positive changes on other stakes identified in this study, such as changing the governance of the data system (with the need to co-build data flows and triggers with local authorities and actors), pushing for data standardization and sharing, and generally speaking encouraging empowerment of actors – in particular locales ones – on data given how critical it is in the anticipatory methodology. However, the voice of local communities has yet to be taken into account when developing data models and trigger values – a topic on which CartONG is currently working.

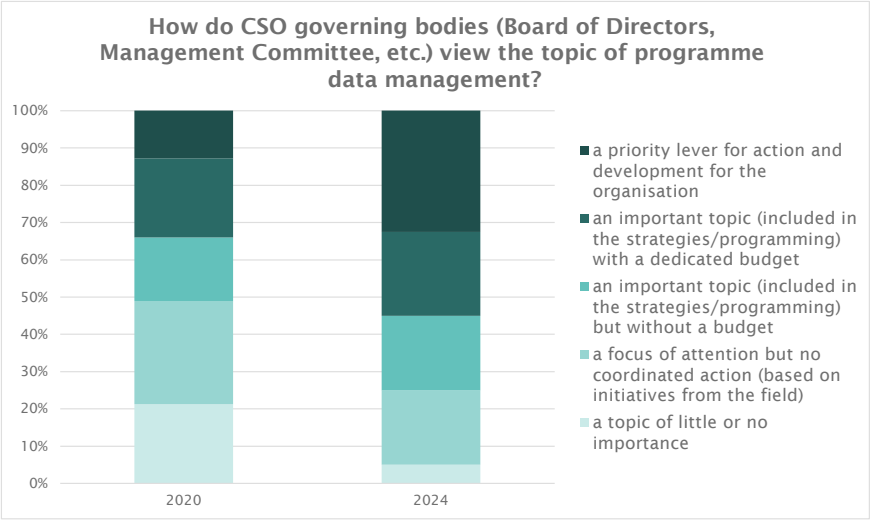
[86] "Forecast based Action by the DREF", IFRC Disaster Relief Emergency Fund, 2019
[87] "Start Fund Anticipation", Start Network
[88] "Anticipatory Humanitarian Action", Matthias Amling, WeltHungerHilfe
[89] "The evidence base on Anticipatory Action", Lena Weingärtner, Tobias Pforr, Emily Wilkinson, WFP, May 2020
[90] "Evidence database", Anticipation Hub
[91] "Does anticipatory action have a role to play in 'wicked crises' like Somalia?", Simon Levine, Lena Weingartner, Alex Humphrey, Muzzamil Abdi Sheikh, Supporting Pastoralism and Agriculture in Recurrent and Protracted Crises Knowledge, March 2023

9.4. CONSISTENT STRATEGIES

We introduced in the first part of our study (chapter 6) the fact that the professionalization of humanitarian data management can only work with a continuum from the individual's skills to the organization's and the sector's policies in general. The latter levels can only be achieved through a resolved and consistent involvement of humanitarian CSO's leadership. As put in our 2022 Human Resource pack on program data, "it is rather tempting to reduce program data management to technical issues, be it the implementation of new tools, the creation of a database or the choice of a case management software. Yet without the (right) people to carry out these functions, and without the right structure to allow them to express their full potential, the task becomes incredibly more difficult"[92].

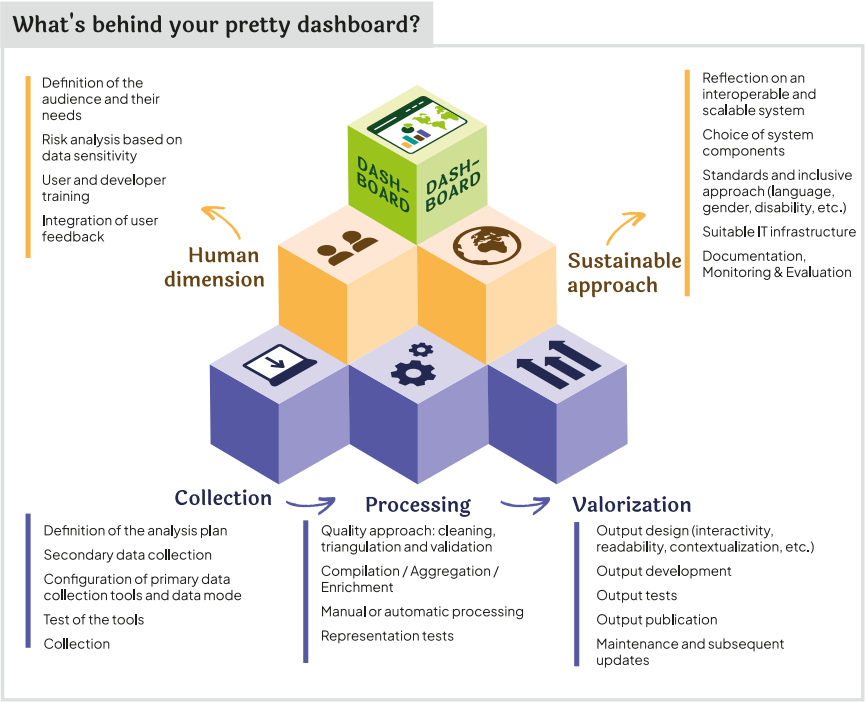
According to our survey's respondents, the endorsement of data management has evolved in the past 4 years, gaining more importance: while only 23% of our 2020 survey respondents assessed that for their leadership it was either a top priority or an important axis, benefitting from dedicated funding, for our 2024 respondents this ratio has more than doubled up to 55%; and reversely the organizations' leadership with no interest in data went down from 20% to 5%:

[92] "HR pack – Program data management for humanitarian aid and international development CSOS – Framing the key issues and getting familiar with the toolbox", op. cit.



One of the reasons behind this evolution is certainly the widespread development of data products (dashboards, maps, etc.) that we have documented through this study: leaders now cannot avoid being confronted both with data tools and with the injunction to be “data-driven”. They however still often underestimate what is required for a data product to be relevant and sustainable (cf. illustration below to see the different prerequisites). Therefore, it seems the battle nowadays for humanitarian data professionals is not so much to create awareness on the importance of data, but rather to sensitize on the need for enough resources to implement it in the right way and for the right reasons. And to push that data is a crosscutting topic useful for programming, advocacy, etc. and not only reporting.

We tried to illustrate in this visual the methodological and technical work required behind the tip-of-the-iceberg map or dashboard seen by the decision-maker. The visual also highlights all the invisible parameters that are needed for data visualization products to be useful over time and also work for the operations and staff that work on the data in a daily basis.



Indeed, in our panel, while a significant proportion of CSOs who responded to the survey mentioned data management as a structuring axis for their organizations with dedicated resources (19%) and/or having a dedicated strategy (11%) or an institutional policy (10%), another large group of CSOs still only have a few tools identified (20%) and/or technical solutions (11%) or procedures (15%) in place for part of their needs. In general international CSOs lie more in the first group than local ones, in particular logically for aspects needing dedicating resources such as putting data management as a structuring axis. A lot of diversity thus remains in the strategic appropriation of data management by CSOs, with operational aspects more advanced due to concrete needs by field teams pushing the topic forward.

The risk of a two-tier system already evoked (cf. chapter 8) is therefore also real for the internal structuring of organizations. While some CSOs (mostly international, mostly big ones) are advancing on their structuring, with dedicated teams, procedures in place or in the making, and building sustainable technical systems and tools, many others (most of the local ones, and most of the CSOs in general except the big ones) still work mostly on an ad-hoc basis. These

organizations use more and more data technologies as documented through the report, but will struggle with building the policies that would allow them to tackle the different stakes defined in this chapter.

> MOST ORGANISATIONS ARE UNABLE TO GIVE THEMSELVES THE TIME TO SIT DOWN WITH THE VARIOUS DEPARTMENTS INVOLVED (IT, M&E, PROGRAM, LOGISTICS...) TO ANALYSE THE INTERNAL DATA ECOSYSTEM, INTERNAL DATA FLOWS/FLOW DIAGRAMS, IDENTIFY A COMMON VOCABULARY AND DATA OR SERVICES TAXONOMIES. WITHOUT THIS, YOU DON'T HAVE AN INTERNAL DATA GOVERNANCE SYSTEM NOR POLICIES (DATA MANAGEMENT, PROTECTION, RESPONSE TO FUNDERS ON DATA SHARING, ETC.), AND WITHOUT THIS YOUR TEAMS CAN'T IMPLEMENT A PROJECT. LACKING A CLEAR GOVERNANCE AND POLICIES, PROJECT TEAMS DON'T KNOW IF THEY'RE GOING TO BE TOLD OFF BY THE MANAGEMENT IN 6 MONTHS' TIME FOR SOMETHING YOU MIGHT HAVE DONE, SO PEOPLE EITHER DON'T TAKE RISKS OR, ON THE CONTRARY, ACT RECKLESSLY – G. COPPI, ACCESS NOW

The sector unfortunately seems to be still driven by risks and crisis more than anticipated efforts: data breaches have pushed the topic of responsible data management to the forefront (cf. chapter 9.6), major cyber-attacks seem to be necessary to create awareness on the level of related threat (cf. chapter 9.7), and on some other stakes leaders might still be waiting for a scandal to see big policy pushed (for instance AI, cf. chapter 9.10; or relations with big tech companies, cf. chapter 9.12). This anticipation is nonetheless the only way to avoid being solution-driven. As an observer of the sector puts it, strategy “helps avoiding situations where a technology is chosen for no other reasons than it being the latest one promoted by a service to one of the directors”.

This picture however needs to be nuanced: the diversity of situations depending on the sector of work have already been mentioned (cf. chapter 6.3), but the same can be said in terms of geographical and cultural diversity. The anglophone humanitarian sector is for instance more prone to a “data-driven” discourse than the francophone one; and in some regions where the level of tech skills, intensity of the international humanitarian response and interest on the topic are hight, such

as the Middle East or Ukraine, data stakes are much more endorsed by local CSOs too.

This representation is also a work in progress as many CSOs have started structuring efforts, proportional to their size, and the landscape could evolve quickly depending on their uptake. We can however only encourage them to continue keeping in mind the strategic and organizational level in their data planning. And advocate collectively for donors also to invest on strengthening systems and organizations, and not only on training or providing tech solutions. A donor in fact mentioned that with the improvement of data they receive from CSOs, they felt less the need to support their capacity building, which shows the importance of advocating for such a comprehensive vision of data management strategies.

To conclude, as a local CSO so concisely put it about mastering data stakes:

> "THIS CAN ONLY BE EVERYONE'S PREOCCUPATION" – DAARA SOS SANTE

9.5. DATA LITERACY

We've just discussed the fact that organizational endorsement of data needs leadership involvement in order to push the entire organization. The method to implement it is data literacy. As synthetized in a blog post CartONG published in 2021, data literacy is identified as a key skill in all organizations, including humanitarian organizations: "data literacy calls for an organisational culture change: creating an environment where data is analysed, criticised, deconstructed, understood, and ultimately used to design relevant responses that advance our mission"[93]. Major humanitarian organizations such as OCHA/Center for Humanitarian data[94], the ICRC[95], or the IFRC have well understood it and developed documentation accordingly[96]. The IFRC data playbook[97] in particular has been for many years a reference in terms of making the topic digestible for

[93] "Why data literacy is important in the aid sector", op. cit.
[94] "Data literacy", Centre for humanitarian data, OCHA
[95] "Analysis and Evidence Strategy 2019–2022 – Better Data, Stronger Analysis, Smarter Decisions", ICRC, 2022
[96] You'll find a list of resources on our Blog: "Top tools and resources to help you start your data literacy journey", Sylvia Musula, IM Portal Blog, CartONG, September 2021
[97] "Data PlayBook", IFRC Solferino Academy

many different types of actors with a learning by doing approach, built with the contribution of a huge number of practitioners from inside and outside the movement.

> THERE SHOULD BE NO SPACE FOR PEOPLE TO SAY "I'M NOT A DATA PERSON" – EXPERT FROM UN OCHA/HDX

Our survey respondents agree with that, placing data literacy as a critical aspect, and commenting on its importance. We have already presented in chapter 6.4 their vision on the level of data maturity of their organization, which is improving but still insufficient in their eyes.

As one French development CSO data specialist testified: "the lack of data literacy in my organization is really an issue, leading each department to set its own procedures (access, sharing, analysis, use…). There is a lot of inefficiency due to the development of solutions in silo, and lot of linked redundancies". Several international CSOs confirmed in our survey this diversity of level between their country operations or sectors. While many mentioned a will to invest in data literacy internally and a general improvement of the situation over time, it seems it's still often individual profiles that drive it, or through learning by doing, with no dedicated approach or policy.

To emphasize this, we can share that a survey held with participants to a 2023 responsible data training organized by CartONG with 250 M&E, IM and program practitioners from 12 humanitarian CSOs also highlighted as number one blocker that data skills were not sufficiently widespread in their organization as their main issues, before any type of technical or methodological expertise[98].

If we take a closer look at dimensions that make data literacy essential to humanitarian CSOs (from the CartONG synthesis):

- Data literacy is critical in a "do no harm approach", as it helps to better understand data and reduce "selection bias" of information. Developing the capacity to interpret the data but also challenge it is essential if we don't want data to reinforce inequalities in our sector.

[98] Responsible data training organised remotely by CartONG with 250 participants from 61 countries & 12 humanitarian NGOs (2023)

- Data literacy helps increase internal collaboration, as it avoids silo approaches where data specialists (information managers, monitoring & evaluation experts, etc.) are the only one managing but also using data. Together with the development of easy-to-understand data products (e.g. maps, dashboards), this allows field staff, who are the producers of data, to make use of it, thus ensuring more engagement and sustainability of the collection process, as well as improving data relevance.

Data literacy dimensions

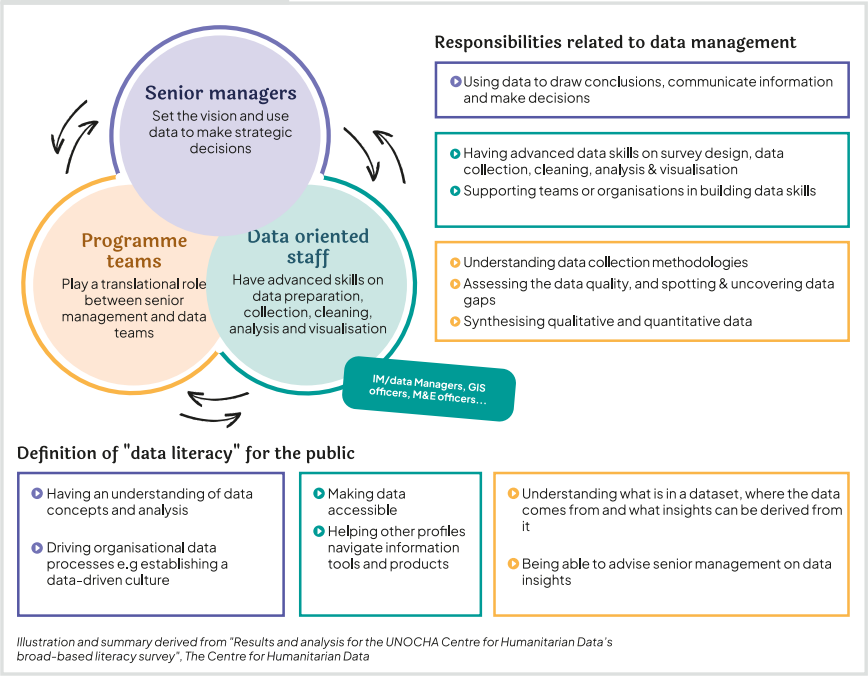


Illustration and summary derived from "Results and analysis for the UNOCHA Centre for Humanitarian Data's broad-based literacy survey", The Centre for Humanitarian Data

- Data literacy increases accountability and transparency internally, with the possibility to question it and avoid creating undue positions of authorities just by owning data skills.
- Data literacy improves organizational performances and efficiency, both in terms of decision-making (cf. also chapter 8.3) but also to cover the different stakes of data management highlighted in this report; it also has impact on future projects, fundraising, operations, HR, etc.

Data literacy is also critical to cybersecurity, that requires the entire organization to enforce protection policies for them to be efficient (this will be covered more in detail in chapter 9.7).

While the benefits of expanded data literacy within organizations and for the sector are clear, most CSOs lack resources for a proper approach towards it. Funding on the topic and buy-in from donors is limited, as it can look like “basic”, “already acquired” skills that appear less innovative than new, fancier tools. As the specialized organization The Engine Room testified, developing a proper data literacy culture is moreover a long-term effort, that also provides lasting benefits (if maintained!) and has an impact both on operational teams and on affected populations.

9.6. RESPONSIBLE DATA MANAGEMENT

Responsible data management could be seen by some actors as something “covered” in the sector, especially since the enforcement of GDPR and the compliance efforts made by many CSOs. However, we’ll see that not only does it still comes up highly in the stakes identified by our respondents, but it also connects to most of the other challenges discussed in this chapter.

The topic of responsible data management, defined as “safe, ethical and effective management of personal and non-personal data for operational response, in accordance with established frameworks for personal data protection” (IASC Operational Guidance on Data Responsibility in Humanitarian Action) has indeed grown considerably over the past few years, in particular for humanitarian CSOs. The pioneering work of organizations such as the ICRC, the Inter-Agency Standing Committee (IASC), the Centre for Humanitarian Data, The Engine Room, or MERL Tech, has allowed many resources to be produced and a general uptake of the topic in the sector[99], including for various specialized sectors[100]. It is now well understood that responsible data management goes beyond “data protection” in consistency with Do No Harm principles – and not just for compliance[101].

[99] Cf. the list of resources identified by CartONG in our Responsible data management Toolbox

[100] Cf. the list of additional sector-specific resources in our Responsible data management Toolbox

[101] Observations from this chapter, besides the elements from our survey and interviews, come from the learning of a responsible data training CartONG organised remotely with 250 participants from 61 countries & 12 humanitarian NGOs (2023)

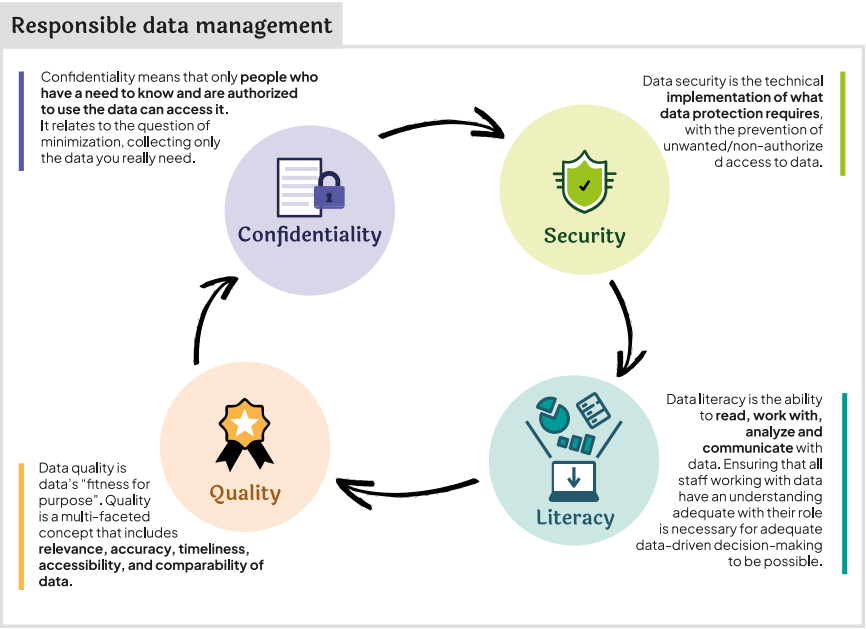
> ENSURING WE 'DO NO HARM' WHILE MAXIMIZING THE BENEFITS OF DATA REQUIRES COLLECTIVE ACTION THAT EXTENDS ACROSS ALL LEVELS OF THE HUMANITARIAN SYSTEM. – IASC OPERATIONAL GUIDANCE ON DATA RESPONSIBILITY IN HUMANITARIAN ACTION

As visible in specialized discussions forum such as the MERL Tech Network, the value of CSOs' data, in particular for the private sector or for unfriendly governments, and the necessity for them to act, has become clearer to them. CSOs are aware they need to protect data that they didn't need to before, and involuntary "data security through obscurity" is not an option anymore. However, while some sectors are – rightly – very concerned and committed on the topic, such as human rights organizations, gaps remain on the appropriation.

This is in particular the case for local CSOs, who in our survey express less interest on the topic than international ones (33% vs. 46%), the topic being less concrete to them as they're less confronted to compliance-linked requests from donors, lack resources to invest such a crosscutting topic, and evolve in a context with sometimes less awareness (absence of national regulation or enforcement of it, lack of specialists of the topic in the country…).

We see it at CartONG as a 4-dimensional vision, including obviously data security and confidentiality, but also with a literacy angle (highlighting the need of good practices throughout organizations: "don't forget that your security is only as strong as your weakest link"[102]) and a quality angle (linking relevance and timeliness of data with decision-making). As we can see, this global vision already connects with other key stakes of CSOs today: cybersecurity (cf. next chapter 9.7), data literacy (9.5), and data-driven decision-making (9.3).

[102] "Responsible data management toolbox – Share and transfer", CartONG, September 2022



This understanding that responsible data management starts with good practices (such as data minimisation or watchfulness on data sharing) rather than high-tech IT security systems is now more and more widely shared. The dynamic continues today: while the uptake of the topic was strongly linked to the enforcement of the European GDPR in 2018, many countries have since developed legislation (protective or repressive), even though there are still important gaps globally[103], that have an impact on the current and future processes and practices of CSOs.

Responsible data is also coming back to the forefront with the recent democratization of AI – such as in the WFP "Palantir" scandal, cf. chapter 9.10 – which is more a revealer of all the responsible data practices that CSOs need to adhere to than a novelty in itself. This is also the se case for the development of digital IDs: wWhile donor support this technology on the contested (cf. chapter 5.1) basis that increases efficiency and reduces fraud, CSOs have little leeway to assess the risk associated with biometrics (while they concern some of the most

[103] Cf. for instance this map by the French Data Protection Commission CNIL

sensitive data that exist) and set up mitigation measures. This technology is also – unsurprisingly – another example where the investment and development of systems is led by the private sector (cf. also chapter 9.12).

The overall evolution of data management – not only for the humanitarian sector – with an ever-increasing quantity of data produced and number of actors intervening on it, means that the question of data security will never fade and on the contrary become always more important. Processes being digitized also means complexification, as it means a double security of digital & non-digital processes.

Responsible data management has become a guiding principle in the approach of many CSOs even though few have (yet!) formalized a dedicated policy. Specialists understand the need to make sense of this need connecting it to humanitarian principles, in particular to onboard field teams that often see these processes as cumbersome or as constraints distracting from the need of the communities they are supporting. The fact that many templates, such as the IASC operational guidelines[104], are now available, also facilitates the embedding in CSOs' strategies and processes (e.g. concepts of data minimization, legal basis for data collection, data sharing agreements, risk analysis, etc.).

Focus: the practice of consent as an example of responsible data practices

Collecting consent has evolved significantly over the past years, going from a tendency of "consent at all cost" from interviewees, to the need for a proper informed consent that gives more power to the interviewees. ICRC[105] and The Engine Room[106] in particular have documented this evolution, and the questioning of the very concept of *informed* consent, in contexts of power imbalance between a CSO providing a service and a potential recipient. The question is therefore whether another legal basis, that then requires a level of data securing that is much higher and therefore more respectful of the rights of the recipients, might not be more adapted in the end[107].

[104] "IASC Operational Guidance on Data Responsibility in Humanitarian Action", Inter Agency Standing Committee, April 2023
[105] "Handbook on Data Protection in Humanitarian Action – Second Edition", ICRC, May 2020
[106] "Unpacking 'informed consent'", Madeleine Maxwell, The Engine Room, November 2019
[107] "Responsible data management toolbox – Consent", CartONG, September 2023

However, hurdles and areas for improvement remain, for which solutions can take many shapes and forms:

Challenge	Possible solution
Need for more accessible terminology	Continuing to translate “data protection” in sector concepts, by acting out the Do No Harm values related to data rather than focusing on the legal and compliance side of things, “showing that it isn’t abstract, it’s about affecting peoples’ lives” (The Engine Room)
Some aspects of the data cycle are still immature in terms of responsible data (data collections built on analysis plans, automated processes for data destruction & archiving, informed consent, data de-identification, AI)	Learning through sector discussions and peer-to-peer exchanges Having technologies in use making user friendly functionalities available
Technologies often evolving too fast for CSOs to keep up adequately	
Level of capacities of operational staff & management often insufficient compared to support staff (in the cases where there are dedicated M&E/IM/data protection focal points)	Importance of taking a transversal approach, embedded with a data literacy one, to raise skills in a concerted way rather than keeping the topic only in the hands of specialists and support functions of ICSOs (cf. chapter 9.5)
Compatible capacities & infrastructures of local partners	
Availability of adequate tools and resources, compatible with field constraints	

Sector coherence and vision: different actors (CSOs, local authorities, governments, donors, UN agencies…) can have a different Do No Harm understanding and agenda	Exploring proactively the risks involved (legal, compliance, reputational…), in particular with real situations that have occurred / scandals in the press
Technology innovation-driven approach often pushed by donors & private tech funders take more and more space, leaving less and less leeway to CSOs (cf. chapter 9.12)	

> IT'S NOT A BUREAUCRATIC PROCESS, BUT IT REQUIRES CRITICAL THINKING, GOOD OPPORTUNITY TO WORK ON HOW POWER DYNAMICS ARE EMBEDDED BETWEEN ACTORS. – BÁRBARA PAES & LESEDI BEWLAY, THE ENGINE ROOM

9.7. CYBERSECURITY

Directly connected to the question of responsible data management (of which we consider it is a component), cybersecurity has emerged as one of the critical challenges of the digital industry throughout all sectors through the past years. It has become a very important component of a responsible data approach, of course regarding data security and data confidentiality, but also connecting to the question of literacy and good digital hygiene practices across the organization.

While it is still seen as a very specialized topic, respondents from our survey rightly identified it as a core stake for their organizations. However, many CSOs can be seen as pessimistic on their ability to face it, assuming that while they can probably withstand small hackers' attempts, they could not counter a state or state-sponsored level attack in any case.

To start with a definition, cybersecurity is according to America's Cyber Defence Agency "the art of protecting networks, devices, and data from unauthorized

access or criminal use and the practice of ensuring confidentiality, integrity, and availability of information"[108].

OCHA's guidance note[109] mentions as potential threats: cyber warfare (from preventing access to basic services to data theft and disinformation), MDH (Misinformation, disinformation and hate speech), cybercrime (cyber-enabled criminal offenses) & cyber sabotage (aiming to disrupt the functioning of ICT). The CyberPeace Institute identifies the different malicious actors as follows: hacktivists (who see world famous organisations as a trophy), criminal groups (for direct ransomware, or through selling to the dark web) or state & state sponsored groups (for organisations whose activities can be seen as disturbing or espionage/sabotage). As described by the Solidarity Action network guidance[110], the biggest threats are phishing attacks (aiming at tricking the recipient into providing sensitive data or installing malware, usually through an email with malicious links or attachments), ransomware attacks, and Distributed Denial-of-Service (DDoS) attacks (involves flooding a network, service or server with excessive traffic to make it cease normal function).

Privacy Affairs (a collective of journalists, experts in cybersecurity and lawyers) had captured that the dark web price for the scan of a passport can go from 50$ to 150$ depending on the nationality looked for, and many of the personal data that CSOs can be storing have an official price on the dark web. CSOs – in particular humanitarian ones working in very critical contexts – indeed hold some of the most vulnerable data in the world, but don't at all have the infrastructure to protect the data.

The high digitalization of CSOs (and of our societies in general, with 70%+ of the world online) over the past decade has led to growing risks of cyber-attacks and increased the digital surface that can be attacked.

The sector, first with the Covid-19 pandemic (for instance WHO noted a fivefold increase in the number of cyber threats during this period[111]) and then the Ukrainian crisis, has been confronted with a new level of attack. In the Ukrainian

[108] "What is Cybersecurity?", America's Cyber Defence Agency, February 2021
[109] "Guidance Note on the Implications of Cyber Threats for Humanitarians", OCHA Centre for Humanitarian Data, April 2023
[110] "Navigating cybersecurity: Guidance for (I)CSO professionals", Solidarity Action Network, April 2022
[111] "WHO reports fivefold increase in cyber attacks, urges vigilance", WHO, April 2020

case, between January 2022 and September 2023, the CyberPeace Institute has documented a total of 2776 cyber incidents conducted by 106 different threat actors[112]. The full-scale invasion by Russian military forces in February 2022 was accompanied by cyberwarfare operations against military and civilian objectives, including humanitarian targets. CSOs are attractive targets (the second most targeted industry by nation-state attacks) because of the sensitive data they handle daily, including political, ethnic, religious, or financial information[113].

Several high-profile attacks on prominent organizations have illustrated this risk:

- The ICRC 2022 cyber hack compromised the data of more than 515,000 people who were part of its Restoring Family Links database, a program created to help people separated by migration and conflict find their family members[114];
- The 2023 NRC cyberattack impacted a project's online database containing the personal information of thousands of project participants, including peoples' names, locations, and contact details[115];
- More recently, the 2024 UNDP cybersecurity incident, from a data-extortion actor[116].

It is worth noting these are situations made public, while it's common knowledge that some situations either are hushed up due to reputational risks or are simply unknown as the CSO may not have the capacity to even know they have been hacked, when the perpetrator's interest isn't directly financial.

Indeed, according to the 2023 CyberPeace Institute NGO analytical report[117], 70% of NGOs either don't think, or aren't sure whether they have an adequate level of resilience to recover from a disruptive cyberattack. And according to the NetHope 2024 State of Humanitarian and Development Cybersecurity Report[118] , nearly two-thirds (65%) of respondents experienced a security breach or critical data incident in the last year, up from 45% the previous year. 85% of NGO think that

[112] "Cyber Dimensions of the Armed Conflict in Ukraine", CyberPeace Institute, September 2023
[113] "Mapping humanitarian tech", op. cit.
[114] "Cyber attack on ICRC: What we know", ICRC, February/June 2022; also presented during GeOnG 2022's Failfest session
[115] "Cyberattack on Norwegian Refugee Council online database", NRC, July 2023
[116] "UNDP Investigates Cyber-Security Incident", UNDP, April 2024
[117] "Cyberpeace analytical report: NGOs serving humanity at risk: cyber threats affecting international Geneva", CyberPeace Institute, 2023
[118] "2024 State of Humanitarian and Development Cybersecurity Report", NetHope, 2024

their staff pose an important risk in terms of cybersecurity, yet only 55 % of them deliver regular cybersecurity awareness sessions[119]. Some big CSOs have indeed started staff sensitising through fake phishing attempts or e-learning modules.

CSOs – especially those with high profile and/or working in contexts where cyber-attacks are frequent – need to understand the threats they are facing, including the risk of possible costs that could go to the level of bankrupting their organization, the complete blocking of their IT systems, or to the targeting of some of the vulnerable populations they're supporting. Exchanges with experts of the sector show that CSOs are now important targets for intrusion and hacking, including because their data processes and bases become more structured, which wasn't the case before.

Based on their experience, the CyberPeace Institute recommends to not consider cybersecurity as purely technical, and that solutions in fact lies mostly in basic good digital practices throughout the organization, building awareness, and putting institutional resources and policies on prevention. Which brings us back again to the question of staff literacy of all things digital, as shown in the following graph.

This of course doesn't prevent the need for investing in strengthening systems: a lot of smaller CSOs still use pirated/makeshift systems, which are particularly vulnerable. The donor funding model, which leaves little margin of core funding for structuring initiatives, is of course limiting the capacity for the long-term investment required. But this is also linked to the fact that the impact of good cybersecurity, or return on investment, is difficult to assess – as success will essentially be defined

[119] "Cyberpeace analytical report: NGOs serving humanity at risk: cyber threats affecting international Geneva", op. cit.

in something (a data breach) *not* happening[120]. Cybersecurity is a continuum that needs to be worked together with partners, which can prove also challenging when governments counterparts (in particular) have limited awareness, skills or tools at hand to contribute (a typical example experience by many humanitarians across the world is government officials using personal WhatsApp or Gmail accounts for data sharing in the absence of functional professional systems, with all the associated risks). On the CSO side, international actors have had the skills and resources to improve their practices, but this is of course less the case of local actors.

Similarly to many of the data stakes discussed in this report (cf. also chapter 9.5), CSOs must walk on thin ice, by confronting the issue while not pushing technical solutions too advanced and inapplicable in field conditions. As discussed recently during a panel at the HNPW conference, CSO's headquarters IT departments sometimes push techno-solutionist approaches, for instance by forcing multi-factor authentication, banning data being saved on external hard drives or imposing strong passwords that need to be changed very regularly[121]. However, in contexts where not every field staff has a smartphone to confirm an authentication, where due to connectivity issues the backing up of data online is not very reliable, or where the skills and time for a password management software to be used are not there, staff will be put in the uncomfortable situation of having to bypass the official secure system with more practical solutions that are much less secure (using another less secure tool, post its for passwords, etc.). It's therefore essential for headquarters to confront themselves to all these field constraints to ensure the solutions they offer are compatible with time, software and hardware in place, as well as with human knowledge and awareness. This will ensure the solutions implemented do not feel like a fish out of water, but rather fit seamlessly into the existing environment.

[120] "2024 State of Humanitarian and Development Cybersecurity Report", op. cit.
[121] Session "Data responsibility: Rethinking 'Do no harm' with a digital lens", CartONG, CyberPeace Institute, International Social Service (ISS) & Access Now, Humanitarian Networks and Partnership Week, May 2024

Focus: useful resources to tackle cybersecurity

For CSOs who don't know how to start on the topic, you can check CartONG's focus on cybersecurity on our Responsible Data Toolbox[122]. It lists several great resources such as OCHA/Center for Humanitarian Data's Guidance note[123], the Solidarity Action Network's Playbook on cybersecurity[124], the Global Cyber Alliance's Cybersecurity Toolkit for Mission-Based organization[125], or the French national information systems security agency's cyber crisis management guide[126].

If we summarise their recommendations, the first steps are to:

- Analyse: evaluate your current infrastructure & the risks you face;
- Invest: have a proactive approach on the topic, allocate resources, get support, think transversal;
- Prepare: have a structured framework, a response plan in case of attack, boost your digital immunity against threats such as viruses & spyware, strengthen the connection to devices and accounts, prevent phishing and malware, diversify supply chains;
- Build capacity widely: concentrate first off on basic digital hygiene, building cybersecurity into everyday working practices.

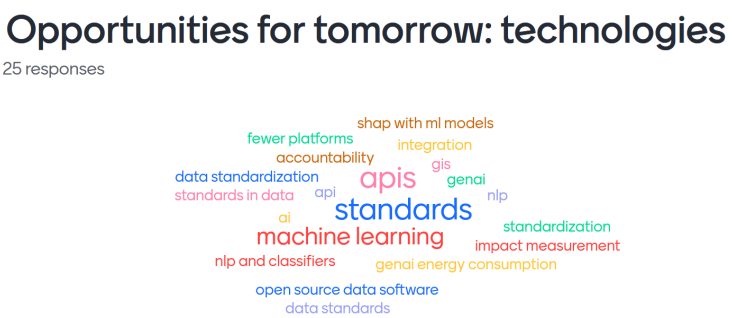
Many of the resources mentioned above can help with each of the steps. It is also possible to get external support, for instance through the CyberPeace institute's builders program[127], a network of corporate volunteers that provide free pre- and post- incident assistance to humanitarian NGOs and make many resources available inside their platform.

[122] "Cybersecurity", Responsible data management toolbox, CartONG, June 2022
[123] "Guidance Note: Data Responsibility and Accountability to affected people in Humanitarian Action", OCHA, August 2023
[124] "Solidarity Playbook on cybersecurity", Solidarity Action Network
[125] "The GCA Cybersecurity Toolkit for Mission-Based Organizations", Global Cyber Alliance
[126] "Crisis of cyber origin, the keys to operational and strategic management", ANSSI, May 2022
[127] "Cyberpeace builders", CyberPeace Institute

9.8. STANDARDIZATION, DATA SHARING & OPEN DATA

Data sharing has been for almost two decades identified as a key challenge for the humanitarian sector. However, following the open data movement in the 2010s, the growth of sharing platforms such as the Humanitarian Data Exchange (HDX) platform (created in 2014), and the widespread development of digital commons such as OpenStreetMap, is there still a need for more data sharing in 2024?

The various sources for this study are clearly advocating that this path has not reached its end – including "external" observers such as donors testifying the difference they see between the humanitarian and international development CSOs, the former having a much lower data sharing culture. When asked what the key technological opportunities were for tomorrow, data specialists from H2H organizations mentioned first standards and APIs (and connected terms such as classification):



These "technical" experts thus identify top solutions rather in the form of partnerships and collaboration than pure technological innovation. Various professionals we interrogated concur that these organizational aspects of doing proper needs assessments, listening to users' needs, identifying the problem to solve before running for a solution, are a key. The Agile method for instance, praised as a solution, too often is still used as a marketing/communication prop rather than transforming project management methods.

Data sharing also has a strong connection to data quality: CartONG's experience over the years (in particular our "data collaborative" projects[128]) has shown that CSOs and even authorities are often reluctant to share data because of this issue. Weakness of methodologies and protocols, lack of metadata and documentation, and insufficient quality control are all factors that push CSOs to not feel comfortable sharing their data. It is in fact similar issues that limit the possibility to use secondary data in the sector (cf. chapter 5.2).

Finally, we will not repeat here the various requirements in terms of data protection that data sharing entails: on top of what was already covered in chapter 9.3, various resources exist on the topic such as guidelines suggested by ICRC and OCHA, which are very much in line with the usual recommendations in terms of responsible use of data[129]. It must be noted that, as other technical requirements, the higher needs in terms of data protection, anonymization, aggregation, etc., that sharing creates, can be a barrier to entry for smaller and local organizations (as even larger ones struggle with data anonymization!).

It is however critical that these solutions remain simple to be widely used, in particular by field actors that are the primary collectors of data (cf. chapter 7). Low-tech basic standards are praised for this by specialists, such as the Humanitarian Exchange Language (HXL)[130], an extremely simplified data standard that can be integrated in any database tool. An organization mentioned the example of simple data sharing through spreadsheet using HXL hashtags for school data in Nepal, which has allowed widespread exchange of data but also their direct use by field actors.

This local, simple approach to data sharing helps with the data equity challenge (cf. chapter 9.1). It can however enter into contradiction with the need from aggregating global datasets. ECHO and the French Development Agency (AFD) confirmed in their interview that the lack of standardization is also an issue for them, one of the difficulties being the capacity to process and aggregate data produced by a multiplicity of project evaluators, each having their own methods. This is why they have different initiatives in place to work on this topic, including

[128] "Data Collaborative on WASH in Malawi & DRC", CartONG, 2018
[129] "Responsible data sharing between humanitarian organizations and donors: towards a common approach", Vincent Cassard (ICRC), Stuart Campo (OCHA), Jonas Belina (Swiss Federal Department of Foreign Affairs), ICRC Humanitarian Law & Policy, June 2023
[130] "HXL Standard", OCHA

for instance AFD funding two initiatives of open data to streamline sharing between programmatic and contextual data. While they are looking for rather high-level data standards to inform their fund allocations and be accountable to their decision-makers, the various constraints they must navigate forces them to enter more details, generating this proliferation of standards and reporting formats without willing it.

As a senior HDX expert pointed out, their strategy is not to define standards or quality metrics: it would be vain – even for OCHA – to pretend to have an in-depth vision on all the sectors' changing data requirements. Their approach is thus, based on the large amounts of data aggregated on their platform, to identify possible quality approaches and promote them to their users, nudging them progressively into building standards. Interestingly, this is very similar to the way the crowdsourcing platform OpenStreetMap builds its data model, through a mix of user field experiences and dialogue to identify and agree on standards.

To conclude, if data sharing is now well accepted as a requirement for humanitarian actors, and in fact is even seen as a visibility asset, there is still much to do in building collaboratively the standards that will allow a streamlined, democratic and general data sharing ecosystem in our sector. And this global standard building effort must always be connected with the need for data equity already discussed above.

9.9. LEGAL AND CONTRACTUAL ENVIRONMENT

The working environment that CSOs currently need to navigate has become a much more complex regulatory one over the past 5 years, between an exponential growth of national legislations, a more structured contractual environment for the different actors and a more and more demanding donor-compliance context. While this question might not look like a data management question at first glance, we'll see why our survey respondents still identified it as a stake for them – for 14% of international CSOs and 19% of local ones.

First, in terms of legislation, the EU General Data Protection Regulation has progressively since 2018 helped set a very good standard that many other legislations across the world use as reference[131]. It took a bit of time for many

[131] "The legal and contractual pillar", Responsible data management Toolbox, CartONG, September 2022

CSOs even for the spirit of such legislation to be respected in related to data on populations affected, but has definitely helped improve data protection practices, with a lot more awareness of the rights of individuals.

A driving trend in African countries around data protection legislation has emerged over the past years[132], with many successful advances in terms of data privacy, such as ensuring that data concerning their citizens should remain in the country. However, there are also often exemption clauses that are too broad, questions on data minimisation or important logistical challenges due to the legislation implemented.

Sometimes foreign legislation that apply to them through donor funding itself raises ethical questions – e.g. the USA "Cloud Act" adopted in 2018 broadened conditions for the US government to request personal data, regardless of the data location. This the case for example if the data is owned by a CSO funded by US Aid, which goes against basic data protection principles and humanitarian values/mandate.

Beyond these cases, over the past few years, a number of countries across the world have set up legislations that – even though presented under the "data protection" lens – can in fact limit the individual freedom of their citizens or give more control to authorities[133].

When local legislation is not in accordance with foreign legislation they need to respect, be it for the "right" reasons or for authoritarian reasons, this leads to CSOs being caught in a legal limbo where they need to be creative in what they respect or not, based also on their mandate and values. This can lead, as shared by various CSO practitioners, to legal, ethical or logistically challenging situations. Some have decided in a given country deliberately to avoid collecting personal data (however problematic in terms of efficiency and accountability to donors) to ensure that they cannot be forced to share it with actors with potentially questionable practices.

The second dimension is the contractual side, that is required when personal or sensitive data sharing or transfers occurs between any type of stakeholders, such

[132] "How governments can strengthen data protection in Africa", Bridget Andere, AccesNow, January 2024
[133] "Data protection laws of the world", DLA Piper, 2024

as donors/UN, CSOs in partnership or consortia, tech companies, local authorities, universities, etc.

This is thankfully much more structured and normalised than previously, with most international CSOs and international organizations now having rigorous data sharing templates and processes. International organizations were 5 or 10 years ago known to regularly benefit from technology providers made available at low cost without protecting sufficiently the data involved, and similarly CSOs were known to share personal data with various stakeholders without questioning the purpose sufficiently…

But the evolving roles of sector actors in relation to tech, with "crossbreeding of roles, with companies taking over humanitarian functions, and humanitarian entities engaging in the direct provision of data and tech services" and related "hybridization" that it entails[134] leads to many new questions that CSOs are not always adequately equipped to tackle.

For instance, the United Nations have their own legal framework, that follows the same general rules as those of the GDPR. Most UN organizations officially have adequate policies and mechanisms in place to adhere to these principles; however local/regional practices are sometimes problematic, leading to unauthorized data transfers to undemocratic States. In other words, UN organisations can have their own legal constraints, political pressure and interests to follow that may in some contexts be contradictory to some of the principles of humanitarian action (see the well-known example of the Rohingya context[135]). Some donors/UN agencies locally can also see CSOs as subcontractors of their own activities, which can lead them to believe that they have a right to the personal data of affected populations they collected. Some therefore continue to propose clauses or annexes in data processor contracts with CSOs that allow them – if signed – to recover all the personal data that the CSOs have collected in the context of the projects (such as the list of affected populations, the Complaints and Response mechanism database, etc.).

Embedded in this question of contractual obligations is the elephant in the room of donor compliance. Accountability to donors and the public is obviously a good thing, putting responsibility to CSOs on their mandate and actions, ensuring they

[134] "Mapping humanitarian tech", op. cit.
[135] "UN shared Rohingya data without informed consent", Human Rights Watch, June 2021

have the proper processes in place, etc. (although this can only be legitimate when donors are also ready to fund the background work it requires, that is long and require proper change management). However, there are also many questionable practices concerning the associated compliance requests from donors in terms of accountability.

The first issue here is the influence of donors on the data collected, i.e. on the definitions of indicators (cf. also chapter 5.1). Indicators and even tools are still routinely imposed to CSOs[136], which forces them to have multiple, often redundant and time-consuming processes that limit their efficiency and adds layers of complexity. 73% of our survey's respondents indicated that the data they collect is still first off used for donor reporting, which is as big a problem as it was in 2020. Respondents also share that, as some donors are catching up with others on this topic, it is becoming even more complex for CSOs, with time consuming and contradictory formats and indicators to capture. CSOs testified that the data requirements are often unrealistic, leading to provide only part of it, with the associated risk of poor evaluation of their performance.

Donors are in fact also sometimes trapped in this system: while they are asked by their supervisors to produce high-level trends to justify the use of taxpayers' money and inform political decisions, a donor executive testified that it's often difficult to them to convert these requests in formats understandable and actionable for CSOs. While his organization tried to align what they needed for accountability and what could also be relevant for the quality of projects and accountability to populations, questions of granularity, quantity and lack of standardization often hampered it, creating misunderstanding among their CSO counterparts (including on the question of data on beneficiaries).

The literature on "over-compliance" is scarce and mostly pointing at financial questions; however, it often connects to quality, M&E and ultimately data. Various interviews we've conducted have pointed out the risk of "data over-compliance", with data increasingly been produced for bottom-up reporting and not used for monitoring, steering of operations, and knowledge building.

A significant example is the screening of beneficiaries: in 2021 the French government issued rules forcing any beneficiary from French public development aid to screen all individuals receiving funds – including the populations benefiting

[136] For instance ActivityInfo for reporting, DHIS2 for health data aggregation, etc.

from the projects of CSOs. Not only did this request go against the responsible data management policies of CSOs, but it also breached humanitarian principles of neutrality. French CSOs and their network head Coordination SUD organized a pushback that went up to the supreme administrative court (*Conseil d'Etat*), and although the guidelines were cancelled in 2023[137]; the situation is still unfolding.

CSOs are questioning the disproportionality of the violation of data protection rules versus the expected results (if the objective is as displayed to fight terrorism). If they are of course willing to abide with compliance requirements that ensure the quality of their work, they also ask to be granted sufficient resources to sustain those different compliance mechanisms (in terms of use of funds but also localization, environmental sustainability, gender, etc.). And not to undermine the rights of affected populations and make their situation even more vulnerable, which not only goes against the humanitarian values of impartiality and non-discrimination, but also leads CSOs to take on a role of identity verification which is not their mandate and fragilizing their relations with populations, jeopardizing the bond of trust that should exist.

Another example if the restricted ability to access some technical tools in some contexts. For instance the limitations in using US-developed tools, and in particular the Microsoft suite, in conflict situations such as Syria, is generating additional costs and delays[138].

On the regulatory environment front, there is unfortunately no easy way to tackle the topic, except through possible experience sharing with peer organisations to share approaches and good practices. Nonetheless, each CSO needs to confront itself to all the legislations that apply for each context of intervention, why not using Data Protection Impact Assessments as a tool.

On the question of compliance and accountability to donors, ensuring there are collective and/or sector discussions or stances on the topic when data is requested by donors or governments without it being justified/legal can help (cf. the example of mobilization of French CSOs mentioned above). Another simple option is

[137] [In French] "Annulation des lignes directrices en matière de criblage par le Conseil d'État", Coordination SUD, February 2023
[138] "Invisible Sanctions: How over-compliance limits humanitarian work on Syria", IMPACT - Civil Society Research and Development, 2020

referring requesting agencies to their HQ data protection policy or general principles, which usually exist.

And more practically, although many of these aspects are more a question of organizational positioning and leadership than of data, vigilance is required to make sure that any data request made by donors goes through a vetting process (ideally an official organizational one, else basing itself on guidance available such as a tool made available by Humanity & Inclusion[139]) to ensure the legitimacy of the request. The objective being that interviewees have accepted it and that the data shared was minimized accordingly to what is necessary (and if possible with de-identification in place, the sharing of a sample rather than full database and other good practices, etc.), to avoid the data being shared by mistake/without realising the consequences by an individual.

Technology in general and data can also be asset for "de-risking" and reducing the compliance burden. A good example is fund transfers: fintech solutions (including blockchain transfers and digital currencies) have allowed to bypass several limitations that the traditional banking system couldn't fix[140]. However, even in this clear-cut example of possible benefits, more research is required to properly assess both the impact and risks of such new solutions.

9.10. Responsible use of AI

Although Artificial Intelligence (AI) in the humanitarian sector emerged as a key topic fairly recently, AI isn't new; it represents a natural evolution for data-driven organizations. However, for those new to this technology, guidance is essential to navigate the transition effectively.

A common misconception is that AI is limited to generative models like ChatGPT. On the contrary, it encompasses a broad spectrum of technologies, including Machine Learning (ML), Deep Learning (DL), Natural Language Processing (NLP), and Computer Vision (CV). Each of these tools offers distinct capabilities that can be applied across various humanitarian contexts and rely of different datasets.

[139] Check "Data sharing agreements", Responsible data management Toolbox, CartONG, September 2023
[140] "Mitigating Financial Sector Derisking through Innovation: The Role of Digital Technologies in Humanitarian Fund Transfers", Dr Erica Moret, NRC, October 2023

A first key aspect to consider is that the development and deployment of AI automation relies on data, and therefore quality data. Indeed, both the training and operational phases of AI require robust, high-quality datasets. However, the humanitarian sector faces challenges with data quality, including missing data points, outdated datasets, biases (cf. chapter 5). American CSOs, that have been spearheading the testing of AI and sector discussions on the topic over the past years (such as with the MERL Tech Natural Language Processing initiative[141]), have realised that getting their data in order for AI to be usable is more of an investment than they initially thought, as shared by Linda Raftree, the founder of the MERL Tech community.

These issues must be addressed to ensure AI systems function correctly and ethically, and not be by-passed by "hard fixes" in data quality (additional remote data processing, with possibly even more AI solutions, rather than improving data skills in the field). For instance, a specialist mentioned attempts to use AI to generate statistics on a country with limited field data collection possibilities based on neighbouring countries, with no methodological guarantee these would be accurate.

> WE AS A SECTOR HAVE SOME BAD DATA HABITS, AND IF WE'RE NOT CAREFUL, THOSE BAD HABITS WILL TRAVEL WITH US WHEN THE SECTOR IS ADOPTING AI AT LARGE SCALE. – A SENIOR EXPERT OF THE HUMANITARIAN DATA ECOSYSTEM

The adoption of AI in the humanitarian sector brings a wide range of ethical concerns to the forefront. There is a notable lack of comprehensive ethical guidance, leading to potential risks of perpetuating existing bad habits. Moreover, the most powerful voices are often overrepresented in AI development, while marginalized groups remain underrepresented, and this can become invisible in AI processes and outputs. This imbalance raises concerns about whose interests AI is and will ultimately serve. AI tools can reinforce existing power structures, potentially centralizing decision-making at the highest levels and leaving affected populations voiceless. There is a growing call for regulatory frameworks at both the UN and national government levels to ensure AI is used responsibly, aligning

[141] "NLP Community of Practice", MERL Tech

with humanitarian principles. A promising first example is UNESCO's Ethics of Artificial Intelligence recommendations[142].

AI offers significant opportunities for the humanitarian sector, such as rapid data analysis (AI can process millions of satellite images in a fraction of the time it would take a human), facilitating accelerated mapping[143], streamlining of back-office operations, making them more efficient and freeing up resources for frontline humanitarian work.

In parallel to its potential, AI presents several challenges:

- **Bias in data**: historical and biased datasets can lead to path dependency, where AI systems perpetuate outdated or irrelevant patterns.
- **Model drift**: as AI models evolve, they may deviate from their original training data, making it difficult to detect when errors occur.
- **Algorithmic errors**: the inevitability of AI errors, combined with the difficulty of detecting them, poses significant risks, especially in high-stakes humanitarian contexts.
- **Ethical considerations**: urgent discussions are needed on how to address issues of ethics, accountability, equity, environmental sustainability and affordability in AI deployment.
- **Community engagement**: there is a misconception that AI is too complex for meaningful community engagement. However, it is possible to involve local populations by explaining AI concepts in accessible ways.

AI is often overhyped, particularly in the context of tools like ChatGPT. While AI can be a powerful tool, it should be viewed as an aid in problem-solving, not a silver bullet. Humanitarians should avoid falling for techno-solutionism – something often happening for ChatGPT, considered as a solution for all problems now – and instead consider AI as a powerful tool to resolve *some* problems and help think through *some* questions.

Furthermore, the sector must remain critical of narratives that paint AI as a universally beneficial solution, especially when driven by powerful firms with vested interests. Several interviewees highlighted that while as for most new technologies "for good"/pro bono initiatives flourish, this shouldn't hide the risk of

[142] "Ethics of Artificial Intelligence – The Recommendation", UNESCO, November 2021
[143] "AI system developed by Stanford researchers identifies buildings damaged by wildfire", StanfordReport, September 2021

aid-washing efforts from big tech companies (cf. also chapter 9.12). In the case of AI, it takes the form of big tech companies opening pro bono programs to feed their learning data with more Global South sources, with the actual purpose of showing they account for diversity; their goal is really to extend their market basis rather than support humanitarian purposes. This could also distort the national publicly-owned data systems by big tech global systems with limited accountability, if governments start relying too much on these technologies. More than ever with AI, the famous saying "if it is free, you (or your data) are the product" must be kept in mind.

The partnership announced between the World Food Programme and Palantir is a well-known example of this risk of sharing data of millions of vulnerable people with a company best known for its work in intelligence and immigration enforcement[144]. While WFP highlighted protective measures such as no personally identifiable data being used, their affirmation to just go for "the best" technical solution without expressing understanding of the associated risks have not reassured data protection specialists. This example also illustrates that even the biggest actors can struggle to influence big tech companies; a donor interviewed for this study confirmed that even for them, enforcing partnership models adapted to the humanitarian sector to big tech companies was difficult if not impossible.

Currently, AI's best applications are towards tedious, repetitive tasks (e.g. back-office efficiency, logistics, image analysis, data processing). In the humanitarian sector, machine learning still has limited use, with the exception of Natural Language Processing (the technology behind ChatGP) and remote sensing analysis[145]. It will however necessarily expand to more strategic applications in a near future.

Language is a good example of the potential opportunities and risks associated with AI: one the one hand, Natural Language Processing technologies offer the promise of unprecedented leap forward in communicating with affected communities in their mother tongue, for instance via chatbots[146]. On the other hand, the power dynamics and source datasets used behind mainstream AI tools

[144] "New UN deal with data mining firm Palantir raises protection concerns", Ben Parker, The New Humanitarian, February 2019
[145] "Mapping humanitarian tech", op. cit.
[146] "CLEAR Tech", Clear Global

can reinforce the exclusion of marginalized groups and minority languages. Indeed, most of the translation tools are for now reserved to a limited subset of languages, mostly from the Global North, where companies developing these technologies have customers.

Several good practices have been highlighted by AI specialists to harness responsibly the power of AI in the humanitarian sector (these recommendations are mostly drawn from the learning programme facilitated by ELRHA as part of their "AI for humanitarians" ongoing initiative[147]):

- **Move slowly and test things**: start by using AI but in a low-risk approach by thinking about the cost of mistakes.
- **Data is at the centre of AI**: good data management is even more important now. AI deployments will be most effective in structured environments where information requirements are clear.
- **Be wary of narratives** of big tech companies pushing for AI, and keep in mind the massive economic stakes for them of proving the positive impact of their tools.
- **Understand how AI makes decisions** because it will make mistakes. It can be very difficult to verify outputs for accuracy and spot algorithmic error and failures. Ensure accountability by having robust checking protocols in place. For example, forcing the model to explain its reasoning process and justify the answers, will not only give better results, but also help to detect errors.
- **Develop AI risk management policies** and data governance strategy as you design your use case, and not necessarily after your pilot project.
- **Stay as problem focused as possible**: start with the problem and develop a list of possible solutions with each solution's benefits and limitations. Then evaluate whether and how AI's use is superior to other approaches.
- **Distinguish between productivity and efficiency gains**: figuring out where those efficiency gains happen geographically, at what level, at what sub national, subregional level is critical. We need to weigh the potential financial benefits against the potentially social negative consequences.
- **Engage with communities**: there are ways to help them understand the concepts behind AI without going into detail on the inner workings of AI.

[147] "AI for Humanitarians: Shaping Future Innovation Learning Journey 2024", ELRHA, December 2023

- **Continue conversations around ethics**: there is still inadequacy at the intersection of data, tech and humanitarian principles.

AI holds great promise for the humanitarian sector but must be approached with caution. By focusing on data quality, ethical considerations, ensuring diverse representation, and engaging local communities, the sector can harness AI's potential while mitigating its risks. As AI continues to evolve, it will be crucial to both engage with the developers of these tools while maintaining a balance between innovation and the fundamental humanitarian principles that guide the sector. CSOs therefore have a responsibility to educate their teams on these stakes, and avoid uncontrolled uses of AI tools, in particular when pushed by the promises of big tech companies (hopefully some networks such as the French coalition of CSOs Coordination SUD have started to undertake this vast effort).

9.11. SUSTAINABLE TECHNOLOGIES

As already mentioned through this report, the humanitarian data ecosystem doesn't live in a desert island. It is therefore essential that it masters the technologies it uses to make them sustainable. Sustainable can have several definitions, but in the context of data we'll understand it as the capacity to be able to use a tool or technology over time. Reversely, the current situation of use of privative tools with no visibility on their availability over time or the sustainability of their environmental and societal impact, cannot be deemed as sustainable.

Giulio Coppi's (AccessNow) recent research on humanitarian tech[148] provides a comprehensive review of the risks associated of dependence to technology. He takes in particular the example of the quasi-monopoly acquired by Microsoft Azure (used by most UN agencies, the ICRC, and many major international CSOs, for instance MercyCorps or NRC), with important risks associated:

> THIS TREND IS MOSTLY GUIDED BY IT-DRIVEN CONSIDERATIONS, AS A SINGLE DIGITAL ENVIRONMENT OFFERS EASIER ACCESS CONTROL MANAGEMENT AND EFFICIENT CYBERSECURITY SETUP, AND BY THE LACK OF REAL COMPETITORS TO THE MICROSOFT OFFICE PACKAGE, WHICH NUDGES COMPANIES TO NEGOTIATE FOR THE FULL BUNDLE OF MICROSOFT SERVICES INSTEAD OF

[148] "Mapping humanitarian tech", op. cit.

> DEALING WITH MULTIPLE AND OPEN LESS-INTEGRATED PROVIDERS. BUT THIS COMES AT A COST. IN ADDITION TO THE RISK OF VENDOR LOCK-IN ALREADY FELT BY MANY INTERVIEWEES, THERE'S ALSO THE SHARED CONCERN ABOUT THE LACK OF OPTIONS BASED ON QUALITY, PERFORMANCE, SAFETY OR ETHICS. THIS REDUCES THE CAPACITY FOR BARGAINING FOR MORE FAVORABLE TERMS OR MORE PROTECTIVE CONDITIONS, MAKING IT ALMOST IMPOSSIBLE TO PUSH BACK AGAINST CONTRACT CLAUSES SUCH AS THE RESERVATION BY THE VENDOR OR ITS PARTNERS TO ACCESS THE DATA TO 'IMPROVE THE SERVICE' OR 'IMPROVE PERFORMANCES'. – G. COPPI

Such dependency has a critical impact on local actors, since they lack resources to be able to switch solutions, and at the same time cannot ensure a proper responsible and ethical use and control of said solutions. But this dependency also impacts larger organizations: for instance a data specialist from a prominent French CSO testified their difficulty to collaborate with other organizations since their infrastructure was based on the Google environment, while most of the organizations in the (French) humanitarian sector chose Microsoft solutions – his organization being for example blocked when trying to use a tool such as PowerBI.

The question of digital autonomy or sovereignty is often connected with the question of environmental sustainability of digital systems. While the possible uses of data tools to fix environmental issues ("data for green") are now well documented, for instance through remote sensing, or via anticipatory action systems (cf. chapter 9.3), humanitarian organizations also need to work on reducing the environmental impact of their tools and systems ("green data")[149].

French CSOs indeed testified that while the topic of infosobriety starts to become a question for data professionals, many practices enter into contradictions and continue promoting (unwillingly) infobesity.

As underlined by the French Environmental agency ADEME, "the low-tech approach involves questioning needs, aiming to keep only the essentials, reducing

[149] Cf. the contents of GeOnG 2022 conference, summarized in: "What was GeOnG 2022 about? A snapshot of 3 days of debate and key learnings", IM Portal Blog, CartONG, November 2022

technological complexity, maintaining what exists rather than replacing it"[150]. This of course doesn't necessarily mean reverting back to pen and paper, but promoting technologies that use less energy and materials, and extend the life of these.

The specialized organization the Engine Room regularly points out the underestimated cost of innovation and technology, CSOs generally ignoring the maintenance costs[151]. Solutions however exist, prioritizing "a low-tech/degrowth mindset that prioritises simplicity and durability over scale and fanciness". The Engine Room mentions several examples: low-tech web design, solar-powered data centres, open source easy-to-maintain code, sustainable AI, etc.

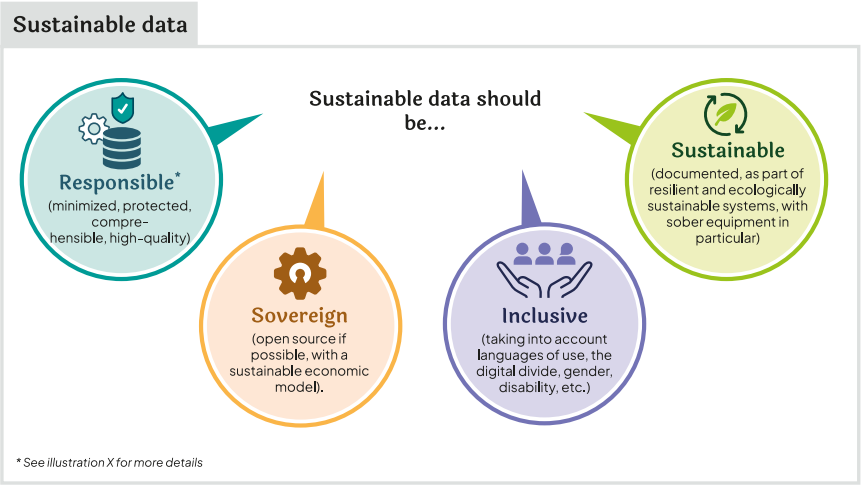
Many CSOs are trying to push in this direction that seem usually more consistent with their mission. This question of digital autonomy (governmental actors would say "sovereignty") still has much less traction than in the public sphere, but it is growing among data and IT professionals, and not only those who are open-source advocates. There is still a long way to go for decision-makers to perceive the strategic aspect of this evolution (as for instance testifies the network of French CSOs Coordination Sud).

One interesting attempt that will need to be monitored in the future is the partnership between the United Nations International Computing Centre (UNICC) and company Canonical (publisher of well-known Linux distribution Ubuntu) to provide to the UN system an open-source cloud, with a focus on security and data sovereignty[152]. Similarly, various CSOs that are particularly committed on either data security and/or digital stakes (Amnesty International, Wikimedia Foundation, etc.) are using the open-source Nextcloud for their could storage.

The use of more sober solutions often offers the side advantage of being more compatible with field conditions. While it might seem an issue from the past for staff at most HQ of CSOs or UN agencies in capitals, broadband connectivity and access to recent devices (computers, smartphones) remain a challenge to many CSOs around the world, in particular local and smaller ones. The same can be said for providing tools that are accessible to non-technical profiles, which are rarer in smaller organizations.

[150] "Démarches "Low Tech", Etat des lieux et perspectives", ADEME, March 2022
[151] "Innovation is not for free", Julia Keseru, The Engine Room, March 2023
[152] "UNICC Partners with Canonical to Build UNICC Cloud", UNICC, October 2023

In general, a consistent and comprehensive vision of "responsible and sustainable data" (linked to "responsible IT") is growingly appearing, which encompasses many of the stakes discussed up to now in this report:



9.12. MASTERED DIGITAL STAKES

To conclude this chapter on stakes for the sector, it seems fitting to cover a challenge that embraces most of the topics we've identified so far, which is our relation to digital technology and innovation in general.

Many specialists who study the use of digital tools by humanitarians make the same conclusions. For instance for The Engine Room, often in their quest for efficiency, CSOs often jump to new tools without measuring the return on investment, diverting resources that could have been used on more basic, sustainable tools[153].

Specialists are unanimous to remind something that could appear a banality but is still a key question:

> FAR TOO MANY PROJECTS SEEK TECHNICAL SOLUTION BEFORE THINKING THROUGH WHAT TYPE OF SOLUTION

[153] "Innovation is not for free", op. cit.

> ARE REQUIRED. WE NEED TO GO BACK TO THE QUESTION: WHAT IS THE PROBLEM THAT THE DATA IS SUPPOSED TO SOLVE? – BILL ANDERSON, DEVELOPMENT INITIATIVES

Or as The Engine Room advises: CSOs need to consider not only the technical lens, but also consider the maintenance of technology (including in human resources), the permanent requirements in terms of data protection and cybersecurity, the imperative to maintain digital inclusiveness.

AccessNow makes the same warning around techno-solutionism as a bypass when humanitarian organizations cannot solve political or organizational problems. In a complex technical ecosystem of cascading providers, and intricate regulations, running after the silver bullet technical solution is a mirage. The example of GeoPoll is revealing: by tracking the various selling and mergers of this company, which carries with it sensitive data about beneficiaries of multiple humanitarian agencies, AccessNow demonstrates the total lack of visibility that most humanitarians face regarding data tools[154]. They take multiple examples of this lack of due diligence in providers that results in serious risks for the data of vulnerable populations, the most jarring one being the collaboration between the United Nations Investigative Team to Promote Accountability for Crimes Committed by Da'esh (UNITAD) with Cellebrite, a notorious surveillance company with track record of collaborating with authoritarian regimes…

Indeed, the humanitarian sector has seen some deep changes in its technical infrastructure that don't seem to have been documented, or even generated by concerted and conscious decisions. For instance, after several years of progressive extension of the Microsoft ecosystems, a reverse movement could be happening with the raise of AI. These changes are often driven by purely technical decisions, with larger actors influencing the smaller ones (local partners in particular) who pursue compatibility with their systems. In fact, even our interviewees from the big players of the sector (donors, major CSOs) mentioned a lack of reflection and sector discussions on these transformations and their impact. Some mentioned also complex influence systems, where decision-making was rather diffuse and not necessarily driven by funders.

[154] "Mapping humanitarian tech", op. cit.

In fact, even at the highest level, the question of power relationships with big tech seems to be eluded. Development Initiatives highlights that the UN Global Digital Compact is built on the assumption that "innovative, interoperable and inclusive mechanisms to enable data to flow with trust within and between countries to mutual benefit, while respecting relevant data protection and privacy safeguards and applicable legal frameworks" can be brokered without conflicts[155]. However, big tech coalitions are already pushing back on attempts at data governance regulation: the Information Technology Industry Council, a coalition of all major big tech companies, advised the US government that "the United States can no longer afford to support nations whose actions are against U.S. techno-economic interests […]This would mean no more aid to countries […] that impose data localization or other digital protectionist tools". While this question of influence of major global corporations concerns more the UN and governments than CSOs, it nonetheless impacts hugely their ecosystem.

Beyond these considerations on the sector's relationship with technology and its providers, raises the question of how innovation is managed by humanitarian organizations – as innovation in the humanitarian sector has been historically focused on technology to drive organisational improvement, with data playing most of the time a key role in it.

Focus: data and humanitarian innovation

In a 10-year review study, ALNAP identifies 4 areas of humanitarian innovation:[156]

[155] "Digital Compacts: Global ideals, regional realities", op. cit.
[156] "Assessing the promise of innovation for improving humanitarian performance: A 10-year review for the State of the Humanitarian System report", op. cit.

It is interesting to read this review with a focus on data.

- For operational innovations, the report mentions several examples of efficiency gains, concluding that "data-driven innovations have been found to improve organisational efficiency, but require agencies to put in place appropriate processes to optimise the use of data and avoid wasted effort" (with a long focus on digital data collection).
- For programme and enabling innovations (for instance on edtech, WASH, or cash & voucher assistance), ALNAP concludes that they "have shown that they can improve the quality, dignity and effectiveness of individual humanitarian programmes. However, there is little research to show the extent to which these innovations have influenced performance at scale".
- For H2H innovations, while recognizing their key innovating role and capacity to address problems not currently addressed in mainstream humanitarian action (for instance the question of language we developed on chapter 5.1), the report insisted to the strong barriers on scaling they face, including not being sector-specific and a business model not fitting in the sector's usual model.
- For local innovations and non-traditional actors, while numerous innovations with high potential have been developed by individual entrepreneurs and start-ups (for instance Dimagi, Ushahidi or Mpesa), the report also concludes on their difficulty to scale and interact with the formal system

For all categories, most of the funding was allocated to prototyping and piloting (71% of total), and only 9% on scaling; and the grants were mostly small (in majority below 100k$, and almost all below 500k$). ALNAP thus concludes on the barriers to scaling innovation, which includes structural issues and lack of incentives (linked to the financing model of the humanitarian sector), difficulty to assess impact (with the lack of possibility to compare), and insufficient overall funding.

The Principles for Digital Development[157], launched in 2015, are now endorsed by more than 300 organizations (including most major UN agencies, a lot of CSOs and support organizations). The principles they promote are of course completely consistent with the different aspects reviewed in this study, but despite this broad

[157] "The Principles for Digital Development", Principles for Digital Development

adoption, remain a challenge. In fact, the 2024's update of the Principles, after a decade of usage, mostly emphasizes a more radical endorsement of the values they were built on: do no harm, inclusivity and local ownership, empowerment, open approaches, dialogue between actors, decision-makers and the private sector[158].

Multiple charters and policies have in fact been developed over the past years to promote responsible and sustainable use of data: Inclusive Data Charter[159], the Data Values project[160], the methodology for the use of human rights-based data of the Office of the United Nations High Commissioner for Human Rights[161], OCHA's Guidance Note on Data Responsibility and Accountability to affected people in Humanitarian Action[162], etc.

CSOs therefore now face the contradiction of on the one hand having to abide with more and more regulations, good practices, guidelines on crosscutting aspects to cover when working with data; and on the other hand, limited when not shrinking resources and injunction to always move towards innovation and new tools. As a big French CSO testified, data reform is often deprioritized in a context of funding cuts (from US or French cooperation notably) which don't leave the possibility to launch crosscutting structuring work. This is in particular the case when they're not seen as priority neither by decision-makers, nor by field operations whose priority is to look for short-term useful solutions. Most of the CSOs who filled our survey made the same alert about the lack of funding (and staff & infrastructure) for the data topic.

> IT'S ALL A KIND OF CONNECTED: THE STAKES OF DIGITAL TECHNOLOGY ARE ENORMOUS, BOTH IN POSITIVE TERMS (BETTER, FASTER DECISION-MAKING, POTENTIAL FOR IMPROVED ACCOUNTABILITY THROUGH CONSTANT IMPROVEMENT OF OUR PRACTICES…) AND IN NEGATIVE TERMS (CYBERSECURITY, INFOBESITY,

[158] "Introducing the Updated Principles for Digital Development", Wayan Vota, ICTWorks, May 2024
[159] "Inclusive Data Charter", Global Partnership for Sustainable Development Data, 2018
[160] "The #DataValues Manifesto: Demanding a fair data future", op. cit.
[161] "A human rights-based approach to data", United Nations Human Rights Office of the High Commissioner, 2018
[162] "Guidance Note: Data Responsibility and Accountability to affected people in Humanitarian Action", op. cit.

> DEHUMANIZATION...). HOWEVER, WHILE CSO ORGANIZATIONAL STRATEGIES PRIORITIZE STRENGTHENING PROGRAM DATA MANAGEMENT SKILLS/PROCESSES TO BETTER MASTER THEM, WHILE PUTTING RESPONSIBILITY (RESPONSIBLE MANAGEMENT…) AT THE CENTER, THIS IS NOT AN EASY TASK FOR ALL ORGANIZATIONS. RAISING AWARENESS AND BUILDING SYSTEMS GOES SLOWER THAN TECHNOLOGICAL ADVANCES. – HQ DATA SPECIALIST FROM AN INTERNATIONAL CSO

Interestingly, another CSO testified on an opposite effect: the growth of awareness on challenges linked to data led to reluctance to digitize processes, instead of trying to find ethical solutions.

To conclude this chapter, CSOs are facing a diversity of challenges and evolutions regarding data management. One of the core question each organization should ask itself is thus its capacity to allocate the necessary resources to onboard correctly new technologies and support innovation. This links to the general question of the management of innovation in the humanitarian sector, which can partly explain why new solutions sometimes fail or do not bring the expected results.

An upcoming study by Groupe URD with Humanitarian Associates exploring the pathways to scale innovations in the sector identified a few success factors: innovation has to be needs-driven to be successful (and not injected by outside innovators); innovation takes time to be adopted, which is often not available in a digital world were obsolescence and competition go fast; evidence is important but perception of innovations matters a lot for their success (communication, marketing, storytelling) and is often less funded. The ALNAP study previously mentioned also assessed that 31% of the innovation – often linked with data – projects they reviewed (from the last decade) were still active[163], and concluded:

[163] "Assessing the promise of innovation for improving humanitarian performance: A 10-year review for the State of the Humanitarian System report", op. cit.

> CONSIDERING THE CHALLENGES, THIS SHOULD BE SEEN AS AN ACHIEVEMENT AND IS COMPARABLE TO THE SUCCESS RATE SEEN IN THE PRIVATE SECTOR. WE CAN'T YET FULLY CAPTURE THE COLLECTIVE IMPACT OF THESE HUMANITARIAN INNOVATIONS. BUT IMAGINE WHAT THEY COULD BE ACHIEVING WITH THE FINANCING AND ORGANISATIONAL INVESTMENTS LEVELS THAT ARE SEEN IN OTHER INDUSTRIES. – ALNAP

10. RECOMMENDATIONS

10.1. ALL ACTORS: PROMOTE COLLECTIVE STRUCTURING, LEARNING AND CRITICAL THINKING

What has been achieved since 2020

- First level structuring of the professional branch
- Better awareness of program data and its stakes, in particular in relation to responsible data

What remains and/or are new recommendations to follow

1. **Evolve towards a more local approach to data**

There are many components to this, detailed in our previously mentioned "Changing the outlook: for a local approach to data" study[164]. In a nutshell they correspond to giving more space to actors in the field to listen to their needs on the topic and answering these needs in an adequate answer (in terms of resources, learning, and funding), making data governance evolve, supporting local data ecosystems and investing in local leaders on such topics. Generally speaking, as the localization agenda is progressing, the data component of it should be seen as an enabler, and not left behind as too complex. This includes of course adapting to the reality of local CSOs, and ultimately leaving them the initiative on defining data and reporting processes.

2. **Continue structuring the professional branch**

Helping all actors have a common frame of understanding, vocabulary and tools to inform their strategies is a key part of the development of the program data branch, that has evolved in the past years but is still in structuration. It will also enable them to be better equipped in their choice of approaches or tools, thus limiting the strategic and technical errors still too often observed.

3. **Have a collective learning approach to program data topics**

Working together, through mutualised resources, training and learning, to better understand and master the topics, approaches and technologies of importance,

[164] "Changing the outlook: for a local approach to data", op. cit.

as individual actors cannot have the leeway to navigate these as well as concentrate on instilling the right level of knowledge and skills to their staff. This includes the sharing and dissemination of resources to help also smaller actors have access to those they need.

4. **Encourage more critical thinking around data creation and usage**

Many questions remain around data that we have seen throughout the study: data not being used for the right purpose, manipulated, or collected inadequately, the absence of data in a remote setting or forgotten crisis giving the impression that there is no need for an intervention, etc.

There is therefore a collective responsibility to more frequently take a step back around data and "think out of the dataset" to learn and grow and avoid questionable situations in terms of data quality or ethics. Data was for a long time associated – through the "data-driven" mantra on everyone's lips – to a purely quantitative vision of M&E, which should continue to evolve more towards an evaluation of change that is wider.

5. **Start by identifying genuine problems before coming out with innovative tech solutions**

Although many new innovative data technologies can be game changers, it's important to always build them based on a pre-identified problem (in particular those identified by local communities), rather than start looking for problems to build a use case for a solution. Technical solutions should remain an enable to achieve

Each tool should remain an amplifier of what a process aims at achieving, rather than an end in itself (which can be encouraged by a higher proportion of "tech innovation-oriented" funding than before).

In the process should always be catered in that the cost of deploying a shiny new technology somewhere is always underestimated, and not just from a financial and resource perspective. It could in fact be bringing more complexity to a territory, could be creating barriers between local and international actors (in terms of capacities, working methods, but also in terms of anchoring the tech in a long-term Nexus approach) and masking underlying problems.

10.2. CSOs: CONTINUE PROACTIVELY INVESTING THE PROGRAM DATA BRANCH

What has been achieved since 2020

- Better integration of program data into programmatic strategies for a number of CSOs
- Ethical dimension of program data made a priority by many organisations
- A higher proportion of organisations with defined roles & responsibilities, and leadership roles covering the topic
- More rationalisation in place in terms of data collection and infobesity
- More specialised staff to take the topic forward, in HQs and in the field

What remains and/or are new recommendations to follow

1. **Integrate program data as a working topic between international and local & national CSOs to ensure the proper uptake of the topic by the latter**

To evolve towards a more local approach to data, it is essential to answer local & national CSOs' needs in terms of capacities and resources, and for CSOs to convene around related topics for exchange and learning. Only thus will local CSOs achieve the necessary uptake on the topic to be more autonomous and proactive in relation to sector needs which will lead to more equitable partnerships. A specific challenge of local CSOs is the relative lack of data experts among their teams, which pleads for even more data literacy efforts among non tech staffs.

2. **Continue building the data literacy level of non tech staff / members of CSOs**

Although data-specialised support staff in the field and in HQ are more and more common across CSOs, the level of data literacy of non-specialised staff is still inadequate. This would however be necessary to ensure that program teams and management positions are in a position to instil their expertise adequately in the data collection exercises that they should be the main users of. It is essential to sensitize them based on their daily needs and constraints (Do no harm values, risks and opportunities related to affected populations but also the organisation itself, etc.). Internal tools and learning formats should be developed to support this capacity building and reduce the gap.

3. **Continue integrating program data into operational strategies in a coherent manner**

Although program data is more and more present in the project cycle, there remains many loopholes, in particular uphill of the data collection and at the end of a project. Prior to the collection phase, it corresponds to ensuring that the relevant data is collected with a proper analysis plan that program teams have been involved it, without it being last minute, or that mixed data collection methods and secondary data are used. At the end of the data cycle, that retention/archiving and deletion is properly.

And beyond the project in itself, making sure that data needs are also thought through programmatically to inform the organisation's wider agenda is an aspect that many organisations still need to improve on (in particular francophone ones). Globally, to always question the purpose of data collection.

4. **Continue building clear leadership and governance on program data**

Program data needs to have a stronger involvement of CSO governance bodies to ensure its implementation is driven with a quality, efficiency and ethical lens, building the internal set-up in terms of roles and responsibilities that it requires.

5. **Explore new topics and technologies proactively like artificial intelligence to be in a capacity to benefit from their possibilities responsibly (for international CSOs mostly)**

Too many CSOs institutionally avoid new topics and technologies until its usage becomes widespread and potentially problematic in the field, if the proper safeguards have not been put in place. It's therefore important to invest the topic uphill to evaluate risks and opportunities rationally, take a stand if required, determine relevant use cases and set up mitigation measures for their usage if necessary. This effort will of course be much simpler if done collectively with other CSOs and pooling resources, in particular through contribution to the different networks and forums that exist.

10.3. NETWORK HEADS: SUPPORTING THEIR MEMBERS IN THE IMPROVEMENT OF THEIR PRACTICES

What has been achieved since 2020

- More awareness of the importance of the topic

What remains and/or are new recommendations to follow

1. **Guide CSOs in the improvement of program data practices**

With their pivotal role in the sector, the uptake of the topic by geographical or thematically-oriented network heads needs to scale up to favour the adoption, by smaller CSOs in particular. It could take the shape of more interaction with the topic and support to their members have a part to play in the consideration of the importance of data issues by CSOs. They must therefore integrate these issues within their team and suggest initiatives and/or services to their members on said issues (training, support, etc.) even if the latter are not directly requesting them. This is particularly necessary for small CSOs that experience more difficulties and need specific support mechanisms.

2. **Establish mechanisms to foster dialogue between CSOs and with support actors**

There are currently limited discussion spaces for CSOs staff that want to exchange and build capacities on data questions, and most are supported by support CSOs (e.g. CartONG's community of practice, Missing Maps and HOT, etc.) or by dedicated secretariat/mechanisms (e.g. MERL Tech, Data4SDGs, etc.). Connecting these existing forums with more traditional exchange spaces of CSOs networks, in particular in the Global South, offers possibility to accelerate the capacity building of the civil society.

3. **Be more proactive to advocate with donors when data-related constraints are too important**

Network heads have a pivotal role in informing donors on the necessary evolutions of funding mechanisms for CSOs. It is therefore key that they not only gather the requirements of their members, but also push proactively ideas (for instance pooled capacity building as mentioned above), not limiting to a "defensive" position (e.g. on questions of compliance).

10.4. Donors and UN agencies: support and encourage the program data uptake of CSOs

What has been achieved since 2020

- More reflexion on the role of program data to support accountability
- Many UN agencies and donor HQs publishing strategies and requirements around data practices

What remains and/or are new recommendations to follow

1. **Ensure the sustainability of pooled and coordinated approach to data resources and training**

To ensure an adequate adoption of the topic by all CSO actors, donors need to think long-term and continue supporting and funding mutualized approaches amongst CSOs and training & learning material that CSOs often find hard to fund (e-learnings, resource platforms, studies, exploration of new topics, etc.). Beyond the production, it is also essential to make these accessible through adequate dissemination, so that they can reach the actors that need them the most.

Training a wider variety of staff of CSOs will help "develop a data mindset" as one donor puts it, showing various staff of CSOs (beyond already convinced specialised staff) what is possible. The experts already can find what the need, the challenge is reaching more people now, going beyond wide but unspecific policies like being "data-driven".

More generally speaking, industry leaders such as UN agencies should enhance the coordination architecture and leadership impacting CSOs. This includes designating focal points for services when relevant and supporting coordination forums. Or, as suggested by a donor, it would make sense to open discussions on the alignment of budget lines for these topics between donors.

2. **Finance and support the capacity building of local CSOs and national data ecosystems**

Finance and support specifically the capacity building (and infrastructure) of local actors on data issues and tools, in particular by letting them define their own strategy and by recognising the priority of their operational needs over the accountability needs of their international partners. As put by Development Initiative, "we should spread data capacity locally to create a bedrock on which a country's digital transformation can be built". This includes CSOs but also investing in national data ecosystems, creating tech communities, encouraging youth to join government initiatives. These ecosystems should then become less dependent on international agendas, which will help avoid data gaps that can lead to underfinancing of certain remote areas or crisis.

3. **Be more proactive in encouraging true accountability to affected populations, by simplifying/reducing expectations concerning accountability to donors**

Donors should open more conversations amongst themselves, with CSOs and with network heads concerning the fact that many CSO see their M&E systems are mostly dedicated to accountability to donors (in terms of tools, indicators, but also processes such as compliance, screening etc). They should work out solutions on how to keep the must-have accountability data and processes, but reducing the CSO workload on such topics if they do not fund them appropriately, by imposing less tools (or building common tools that are relevant for CSOs), standardizing some indicators while accepting a more local and contextualized approach, including more qualitative research methods. This would help CSOs, in particular smaller ones, dedicate proportionally more of their limited program data resources to accountability to affected populations, which should play a more central role in accountability to donors that it does today.

4. **Encourage conversations around data governance**

The role of data in current localization debates is more and more obvious. Donors could support discussions on such topics, highlighting the importance of data governance as well as systems that empower affected populations, seeking ways to better protect human rights in digital spaces.

This includes also building on local activism and community-based leadership for effective response, challenging data divides, in which a small handful of actors control and benefit from enormous amounts of data, to reshape power structures. And strengthen collaboration with governments, to streamline the links between national systems and humanitarian ones, respecting each parties' mandate and policies.

5. **Build balanced relations with Big Tech/private providers**

Rather than – sometimes relatively blindly – partnering with the private sector, that may have different long-term agendas than humanitarian actors, it is important that donors also open discussions with them and challenge them if needed to ensure the tech in question is compatible with sector values in the long run. This is particularly the case for AI tools, to ensure new technologies improve the efficiency of the sector without compromising humanitarian principles, taking into account the risk of disinformation, data breaches and bias.

6. **Fund impact evaluations related to program data to inform future funding and encourage the scale up of innovation funding**

Donors should experiment funding impact evaluations around data activities and capacity building. This will create the conditions for increased trust and transparency of innovation funding, and hopefully a scale up that would allow the humanitarian industry to reach similar levels of innovation funding than international development, or even the lowest-spending industries in the private sector (such as paper or basic metal industry – which invest more than the humanitarian sector on innovation…)[165].

10.5. For specialised Support CSOs: better assist CSOs navigate the data stakes of the day

What has been achieved since 2020

- The network of (data) support CSOs continue to grow and structures, in particular through the H2H Network, with more capacity to advise and influence the ecosystem
- This leads to support and services that are more attuned to the needs of CSOs and helping them navigate new stakes (localisation, accountability to affected populations, cybersecurity, etc.), although some topics strongly related to data still have gaps (for instance IT)

What remains and/or are new recommendations to follow

1. **Continue the reflection and advocacy on larger stakes and evolution of practices while developing products for the sector, in a collaborative approach**

Support CSOs have a social mandate that should encourage them to cater into their work and posture with various sector actors' aspects such as sustainability, localisation and a more responsible approach, to help bring about the changes that are needed with a data angle, beyond only focusing on answering immediate needs with innovative tech. Support CSOs should leverage their collective

[165] "Humanitarian R&D Imperative - How other sectors overcame impediments to innovation", Deloitte/World Humanitarian Summit, May 2015

capacity by continuing investing in their networks and collective advocacy, including but not limited to funding aspects.

2. **Continue supporting an improved quality & relevance of data, but keeping it proportionate to the needs**

Support CSOs often have the capacity to have a very strong understanding and skillset concerning all the components of data quality and responsible data practices. It is therefore important that they be a driving force with their various partners to take the topic up as required, consider the level of quality that is required compared to the associated investment, and also question and highlight data gaps that they might be aware of proactively.

3. **Continue exploring how the assessment of program data can be made**

Many donors finding it hard to justify the direct funding of program data activities as part of project implementation due to it being a “black box”. Making their case by doing what is possible to evaluate the impact of such activities (even though it can be much harder than for traditional activities) would be useful. Support CSOs are well placed for such exercises, analysing the benefits and constraints of such activities to bring more tangible elements to sector and donor discussions on the topic.

11. CONCLUSION

This overall panorama could create a feeling of being overwhelmed for CSO data practitioners when reading it fully (if you just did so, thanks!). The challenges are numerous, CSOs are not necessarily in the driving seat to face them, and the situation keeps evolving fast.

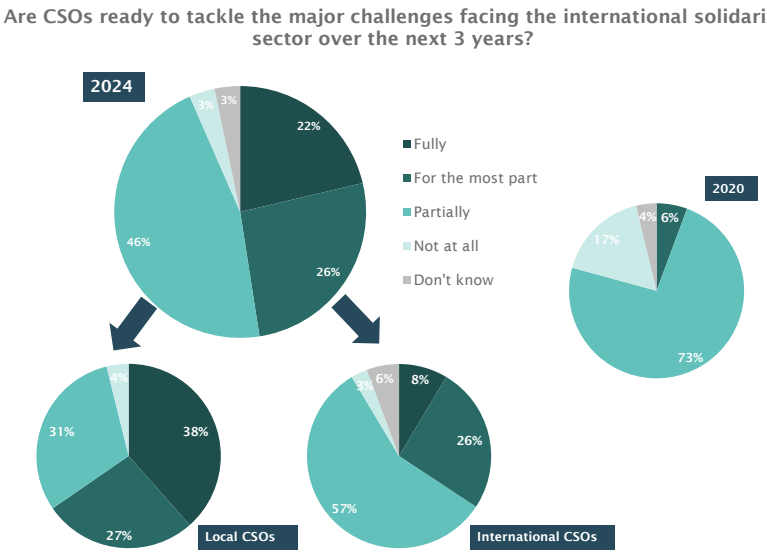
Nevertheless, we have also highlighted all through our study many examples of impactful case studies, positive policies' evolutions, and successful collective endeavours to transform the system. We have thus built our reflections, and the recommendations issued from them, to try to help CSOs navigate the complex sea of data in the coming years.

To conclude this exploration, we'd like to leave the floor to our survey respondents, who not only have dedicated significant time to answer our questions, but also have shown how serious humanitarian practitioners take the stakes around data.

> WE HAVE DEVELOPED A DATA STRATEGY TO GUIDE DATA GOVERNANCE, SECURITY, QUALITY, ARCHITECTURE AN CULTURE IN COMING YEARS. HOWEVER, CHALLENGES WILL DYNAMICALLY EVOLVE AND SO DO WE - WE ARE WORKING ON BEING READY FOR THEM! – WELTHUNGERHILFE

Our survey respondents are in fact rather optimistic on their ability to face these multiple challenges. They estimated the readiness of their CSO much higher than when we produced our previous version of the study in 2020. In fact, 47% of the respondents of the survey felt their organization was "completely" or "largely" on top of the stakes discussed in this study – a figure up from the 6% of our respondents in 2020. Only 5% of respondents felt either their organization was totally unarmed facing these challenges… or not concerned!

Intriguingly, this figure was even higher (63%, including 37% of "completely") for local CSOs. Since few of them gave more details, it is hard to know if that they understood this question as "being aware of the stakes" rather than "having the capacity to tackle them".



Anyway, this general feeling of improvement and optimism by CSOs professionals puts in perspective the many challenges developed through this study, and show humanitarian are willing to face them. We can only hope this work will help them!

CartONG

23 Boulevard du Musée
73000 Chambéry (France)
+33 (0) 4 79 26 28 82

info@cartong.org
www.cartong.org