IFRC

INVESTIGATING SAFE DATA SHARING AND SYSTEMS INTEROPERABILITY IN HUMANITARIAN CASH ASSISTANCE

ACKNOWLEDGEMENTS

Funded by European Union Humanitarian Aid

This research and publication were made possible thanks to funding from the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

This report was written by Caribou Digital, led by authors Bryan Pon, Alphoncina Lyamuya, Emrys Schoemaker, and Horacio Nunez. We extend our sincere gratitude to all the interviewees and reviewers for their generous contributions.

Caribou Digital

Thanks also to the members of the Dignified Identities in Cash Assistance consortium (DIGID), the Norwegian Refugee Council, the Norwegian Red Cross and Save the Children Norway for their involvement and support during this research.

NRC NORWEGIAN REFUGEE COUNCIL

Norwegian Red Cross

Save the Children

Contact us:
Requests for commercial reproduction should be directed to the IFRC Secretariat:

Address: Chemin des Crêts 17, Petit-Saconnex, 1209 Geneva, Switzerland
Postal address: P.O. Box 303, 1211 Geneva 19, Switzerland
T +41 (0)22 730 42 22 | **F** +41 (0)22 730 42 00 | **E** secretariat@ifrc.org | **W** ifrc.org

American Red Cross disaster staff and Philippine Red Cross volunteer review the distribution plan before leaving a base camp in Tacloban, Philippines. (Photo Credit: American Red Cross)

> *We envisage solutions whereby interoperable, non-proprietary, data registries can allow a level of data sharing between humanitarian agencies and private sector service providers that is safe, secure and improves humanitarian programming through enhanced accountability. Data protection safeguards need to be in place, and people should be consulted on the use of their data.*
>
> *– Donor Cash Forum Guiding Principles on Interoperability of Data Systems in Humanitarian Cash Programming.1*

Salvadoran Red Cross supports families from communities with the humanitarian cash transfer assistance program amidst COVID19. (Photo Credit: Cruz Roja Salvadoreña)

1. Donor Cash Forum (DCF). Donor Cash Forum Statement and Guiding Principles on Interoperability of Data Systems in Humanitarian Cash Programming www.calpnetwork.org/wp-content/uploads/ninja-forms/2/DCF-Interoperability-Statement-FINAL.pdf

CONTENTS

Abbreviations	3
Executive summary	4
Project overview	5
Methodology	6
About technical interoperability	8
Current approaches	14
Alternative approaches	21
Lessons from other industries	23
Evaluation	25
Recommendations	30
Conclusions	34
Bibliography	36
Appendix A. Data-sharing template examples	38

ABBREVIATIONS

API	Application programming interface
CCD	Collaborative Cash Delivery Network
CVA	Cash and voucher assistance
CWG	Cash working group
DCF	Donor Cash Forum
DIGID	Dignified Identities in Cash Assistance
ECHO	European Civil Protection and Humanitarian Aid Operations
FHIR	Fast Healthcare Interoperability Resources
GDPR	General Data Protection Regulation
ID	Identity
IFRC	International Federation of Red Cross and Red Crescent Societies
OCHA	Office for the Coordination of Humanitarian Affairs
PRIMES	Population Registration and Identity Management Eco-System
UNHCR	UN High Commissioner for Refugees
WFP	World Food Programme

EXECUTIVE SUMMARY

The simplest, most common way to share data across humanitarian organizations remains sending a spreadsheet in an email. This approach uses tools and technical knowledge that are low cost and nearly ubiquitous; it is the lowest-common denominator approach that almost every organization can use to collaborate with other organizations.

On the other end of the spectrum, the largest and most well-resourced humanitarian organizations have built sophisticated software systems with global databases, robust APIs, biometrics capabilities, and even distributed ledger technology (i.e. blockchains). The question of interoperability in the sector is therefore in part a question of how to reconcile these wildly disparate systems, organizational capacities and governance structures, while maintaining or improving on security and data protection measures.

The interest in improved interoperability is driven by a growing recognition that as humanitarian aid is increasingly digitalized and data driven, the need for improved coordination across organizations and technical systems has become core to the effective delivery of services. This is especially true with cash programming, which offers tremendous scale and leverage when using digital channels, but also requires even more coordination and sharing of data with a broader range of stakeholders, including private-sector financial institutions.

This report offers a technical landscaping of the current state of data sharing and interoperability in the sector, and evaluates the potential of emerging technologies and operating models to address challenges in current approaches. The analysis is focused on cash programming, and specifically on two use cases: deduplication of beneficiaries, and referrals of individuals.

Key insights from this workstream:

- **There is no silver bullet technical solution** – Given the tremendous diversity in organizational capacity and resources, operating environments, data protection regimes, stakeholder relationships, and nature of the response, there is no singular solution or technical approach that is ideal in all contexts, even for a single use case. Instead, we believe a plurality of technical approaches will continue to be used for data sharing, and that this plurality is healthy for the sector overall – e.g. it reflects the demands and capacity of local context, and provides resilience against monopolization and failure.
- **Spreadsheets still matter** – While there have been significant advancements in technical solutions for data sharing – in particular, for deduplication – a huge number of organizations still rely on emailing spreadsheets and other insecure methods. Many if not most recognize the risks associated with processing and sharing data in these ways, but the lack of resources and technical capacity limit their ability to procure more robust systems.
- **Interoperability is not just a technical challenge** – Although this analysis is technical in nature, it's important to recognize that technical systems are just one aspect of interoperability. The 'interoperability stack' model (see section 'About technical interoperability' below) highlights the need for alignment at the organizational and legal layers as well; in fact, these non-technical aspects are likely more challenging than the technical layer.[2]
- **Power asymmetries shape data sharing** – There are significant power asymmetries between the UN agencies and international NGOs and smaller, lower-resourced NGOs, and these asymmetries are reflected in the access to and control over data-sharing technologies; in short, larger organizations have systems, power and control, and smaller organizations typically have little agency in how they engage with technical data-sharing processes.
- **Data standardization is the key** – Regardless of the solution or technical approach, effective data sharing requires alignment on data standards, which we see as a clear opportunity for improving the foundation and potential for interoperability across the sector. Importantly, data standardization is not threatening to any actors, and in fact was explicitly encouraged by virtually all key informants, including private sector vendors.

2. Recognizing this, the complementary project led by the Collaborative Cash Delivery Network (CCD) is focused on developing new approaches to the governance layer as a key way to unlock more collaborative and effective data sharing.

Project overview

The Viet Nam Red Cross Society distributed Multipurpose Cash Assistance targeting 130,000 people affected by extensive flooding due to heavy monsoon rains in the Da Teh district in Viet Nam. Photo Credit: IFRC)

This project on interoperability[3] is led by the Dignified Identities in Cash Assistance (DIGID) consortium[4] and is funded by the European Civil Protection and Humanitarian Aid Operations (ECHO) as an effort to strengthen the humanitarian sector's ability to securely share data related to the identification of beneficiaries and services provided to them.

The project has four phases: 1) landscaping to understand the challenges in data sharing and interoperability, 2) evaluation of current and alternative technical approaches to data sharing, 3) a technical validation using a prototype or simulated environment, and 4) developing a roadmap for how progress on interoperability may continue past the project period. The findings of Phase 1 have been captured in a landscape mapping report, along with detailed analyses of four different data-sharing use cases.[5] This report summarizes the findings of Phase 2.

Running parallel to this DIGID project, the Collaborative Cash Delivery Network (CCD) is also working on ways to advance interoperability, with a focus on governance frameworks, digital literacy and the theme of data stewardship.[6] The two projects will be coordinating their efforts in an in-country testing phase, where the complementary workstreams – one more technical, one more governance oriented – can jointly serve to validate the findings of the research through activities with local stakeholders.

And finally, note that the scope of this project is limited to the systems and processes of humanitarian organizations and other institutional actors. It does not foreground the experiences and needs of individual beneficiaries and other affected persons, not because these are not paramount considerations in how humanitarian organizations fulfil their missions, but simply to provide a focused technical and operational analysis.

3 For more details, see the project page at https://interoperability.ifrc.org

4 DIGID is a group of four NGOs: the International Federation of Red Cross and Red Crescent Societies (IFRC), the Norwegian Refugee Council, the Norwegian Red Cross and Save the Children Norway. The four organizations previously collaborated on the DIGID project with the objective of providing a digital identity (ID) for people without official ID so they can access cash assistance. See more at https://hiplatform.org/digid

5 Access all reports from Phase 1 at https://interoperability.ifrc.org/services-1

6 Access information on the CCD project at www.collaborativecash.org/copy-of-thematic-focus

Methodology

Joyce Eldoret receives cash assistance as part of the drought response through Mpesa mobile money in Lokori, Turkana county, Kenya (Photo Credit: IFRC)

The research approach employed for this workstream was a combination of literature review and in-depth interviews with key informants representing different roles and perspectives in the sector. We interviewed 28 individuals across the following categories:

- **Humanitarian technical operators** – Those who have direct experience using technical systems to share data in cash programming contexts. For these individuals, we sought to understand their current systems, workflows and pain points[7] in regards to beneficiary management and data sharing for cash programming, and tried to validate our mental models for how different components of a technology stack play distinct roles in enabling data-sharing practices.

- **Humanitarian technology vendors** – Organizations developing systems that play a direct or indirect role in data sharing among humanitarian organizations, e.g. KoboToolbox, RedRose, Building Blocks, Population Registration and Identity Management Eco-System (PRIMES). For these individuals, we sought to understand the role that their solutions play in cash programming information management, how they currently integrate or connect with other systems, and the extent to which they believe their systems can support different forms of interoperability.

- **Humanitarian leadership** – Those in non-technical roles at international NGOs and UN agencies who are managing cash and voucher assistance (CVA) programming, participating in cash working groups (CWGs) and similar bodies, negotiating contractual agreements, or those involved in CVA in other ways. For these individuals, based on their role, we sought to understand the organizational perspective on barriers to data sharing and interoperability,

7 'Pain points' refer to areas that are time-consuming, introduce risk, are expensive or affect the dignity of affected populations.

how the organization makes decisions around technical investments and procurement, technical capacity and resourcing constraints, and similar topics

- **Humanitarian associations** – Industry associations and consortia that are organized around cash programming, standards or data protection issues. For these individuals, depending on the association mission and role, we sought to understand their perspective on the key barriers and potential levers for improving interoperability.
- **Non-humanitarian technical operators** – Organizations and experts outside the sector who are involved in technologies and standards that enable different forms of data sharing or interoperability. For these individuals, we sought more generalist perspectives on how current and emerging technologies are addressing the fundamental challenges – from a technical point of view – around the sharing of personal information.

The Nepal Red Cross unconditional cash assistance in Balaju, Kathmandu. (Photo Credit: American Red Cross)

About technical interoperability

Romanian Red Cross volunteer, Svitlana Rostetska, works with two refugees to help them enroll for financial assistance at a service point in Bucharest. (Photo Credit: IFRC)

From healthcare to government agencies, the ability to share and integrate information is critical for making informed decisions and driving innovation. In the humanitarian sector, interoperability can be understood as the ability of different organizations and systems to collaborate through the sharing and analysis of data and resources to facilitate operational processes and meet the needs of affected populations.[8]

From a technical standpoint, interoperability refers to "the ability to access and process machine-readable data from multiple sources, sometimes automatically, without that data losing meaning or integrity."[9] Taking a broader view, the Donor Cash Forum (DCF) uses a definition from ECHO[10] when referring to interoperability as "the ability of organizations to interact towards mutually beneficial goals, involving the sharing of information and knowledge between organizations, through the business processes they support, by means of exchanging data with other systems using common standards."[11]

There are different aspects to interoperability, reflecting the fact that interoperability is more than simply connecting different datasets and systems and enabling information to be shared. It is a socio-technical phenomenon, characterized by elements that include legal, organizational, semantic and technical aspects.[12] Achieving and maintaining interoperability is also more than a one-off effort –

8 IFRC. 2023. Enabling Dignified Humanitarian Assistance Through Safe Data Sharing: Landscape Mapping. https://interoperability.ifrc.org/services-1; Currion, P. 2022. Data Portability and Digital Identity in Humanitarian Aid: A Desk Review; www.collaborativecash.org/data-portabilitydeskreview; Data Interoperability Collaborative. 2017. Improving Data Interoperability for the SDGs (meeting).

9 Development Initiatives. (2017). The frontiers of data interoperability for sustainable development. https://devinit.org/resources/data-interoperability-sustainable-development

10 DG ECHO. 2022. Thematic Policy Document No 3. Cash Transfers. Directorate-General for European Civil Protection and Humanitarian Aid Operations. https://ec.europa.eu/echo/files/policies/sectoral/thematic_policy_document_no_3_cash_transfers_en.pdf

11 DCF. Donor Cash Forum Statement and Guiding Principles on Interoperability of Data Systems in Humanitarian Cash Programming www.calpnetwork.org/wp-content/uploads/ninja-forms/2/DCF-Interoperability-Statement-FINAL.pdf

12 Allen, D. K., Karanasios, S. & Norman, A. 2014. Information sharing and interoperability: The case of major incident management. European Journal of Information Systems, 23(4), 418–432. https://doi.org/10.1057/ejis.2013.8

it is an ongoing effort and something that emerges through building trust and relationships between individuals and entities working to achieve common objectives.[13]

The focus on achieving interoperability in the humanitarian sector is largely catalysed by the growing use of digital tools and platforms by humanitarian organizations to streamline operational responses, including the coordination and administration of CVA to communities affected by humanitarian crises across the world. With the potential to improve the quality of humanitarian responses, enhance beneficiaries' experiences, and promote effective coordination and collaboration among stakeholders in aid delivery ecosystems, the need for interoperability is more critical than ever.[14]

The landscape mapping report from the first phase of this project highlighted that interoperability efforts have the potential to counter duplication of cash-based assistance; promote effectiveness, efficiency and quality of humanitarian service delivery; bolster the vision for moving towards a more people-centred approach to humanitarian assistance; and reduce risks and promote enforcement of data protection.[15]

Technical functions are critical (though insufficient) elements to enabling interoperability, and commonly include both data standards[16] and technical interfaces like application programming interfaces (APIs) as crucial enablers for technical interoperability.[17] APIs can "allow streamlined access to a defined set of data or functionality, sometimes combined with an authentication function to ensure the user has granted their consent".[18] This not only enables the transfer of data (such as personal identifiable information) across systems but also supports other goals of interoperability including data privacy and security. APIs also help to define ways various software components interact and formats through which data can be shared between organizations or bodies.[19] Despite APIs' ability to facilitate data accessibility and system interoperability, "they may not on their own guarantee the seamless use of data from one platform by another".[20] Other elements beyond the technical are thus required for truly operational interoperability between systems.

This workstream sought to map potential technical approaches to interoperability against current operating and technical contexts. Focusing on the semantic and technical components required for interoperability between systems, this phase examines effective approaches central to addressing challenging aspects of data sharing among humanitarian organizations, particularly around deduplication and individual referral use cases.

Supporting environments for interoperability

Practical and adaptable data-sharing principles and mechanisms are central to improving system interoperability, particularly given ramped-up efforts to digitalize cash and voucher programming. O'Brien (2017) emphasized that it is crucial for the right data infrastructure to be established as a mechanism to ensure robust integration of existing "fragmentally collected data within a secured network." And to ensure the effective use of technology capabilities, OCHA (2021) suggested the need to "build and promote interoperable data-sharing platforms with adequate protections for personal and sensitive data within and across organizations and sectors".[22]

13 Mistry, P., Maguire, D., Chikwira L. & Lindsay, T. 2022. Interoperability is more than technology. The King's Fund. www.kingsfund.org.uk/publications/digital-interoperability-technology

14 IFRC. 2023. Enabling Dignified Humanitarian Assistance Through Safe Data Sharing: Landscape Mapping. https://interoperability.ifrc.org/services-1; Madon, S. & Schoemaker, E. 2021. Digital identity as a platform for improving refugee management. Information Systems Journal, 31(6), 929–953. https://doi.org/10.1111/isj.12353

15 IFRC. 2023 (Ibid).

16 Development Initiatives and Publish What You Fund. 2017. The frontiers of data interoperability for sustainable development. www.publishwhatyoufund.org/app/uploads/2017/11/JUDS_Report_Web_061117.pdf

17 See also OECD. 2021. Data Portability, Interoperability and Digital Platform Competition. www.oecd.org/daf/competition/data-portability-interoperability-and-digital-platform-competition-2021.pdf; Riley, C. 2020. Unpacking interoperability in competition. Journal of Cyber Policy, 5(1), 94–106. https://doi.org/10.1080/23738871.2020.1740754; Brown, I. 2020. Interoperability as a Tool for Competition Regulation. Open Forum Academy. https://openforumeurope.org/publications/ofa-research-paper-interoperability-as-a-tool-for-competition-regulation

18 OECD. 2021. (Ibid.), p. 12.

19 Gal, M. S. & Rubinfeld, D. L. 2019. Data Standardization. New York University Law Review, 94, p. 750.

20 OECD. 2021 (see above), p. 13; Gal, M. S. & Rubinfeld, D. L. 2019. (Ibid.)

21 O'Brien, S. 2017. This is how we build a stronger, data-driven humanitarian sector. World Economic Forum. 13 January 2017. www.weforum.org/agenda/2017/01/this-is-how-we-build-a-stronger-data-driver-humanitarian%20sector

22 OCHA Centre for Humanitarian Data. 2021. OCHA Data Responsibility Guidelines 2021. https://centre.humdata.org/the-ocha-data-responsibility-guidelines, p. 54.

But both data sharing and system interoperability require responsible use, exchange and management of information and data about the vulnerable populations humanitarian that organizations intend to serve. The OCHA Centre for Humanitarian Data defines data responsibility in humanitarian action as "the safe, ethical, and effective management of personal and non-personal data for operational response".[23]

To hold themselves accountable and ensure both parties are clear about their roles and responsibilities around data sharing in humanitarian settings, organizations such as UN agencies and international NGOs in the cash delivery ecosystem use data-sharing agreements. For instance, to realize the benefits that can be generated from interoperability in cash-based assistance and developmental goals, multiple UN agencies (UN Children's Fund (UNICEF), UN Office for the Coordination of Humanitarian Affairs (OCHA), UN High Commissioner for Refugees (UNHCR) and the World Food Programme (WFP)) came together in 2018 to sign the UN Common Cash Statement, a multilateral agreement that includes data interoperability and systems development as a core commitment.[24]

This complements bilateral data sharing agreements and collaboration between UNHCR and WFP specifically to better integrate their data systems,[25] and importantly, the definition of a minimum core assistance delivery dataset that defines specific data fields and standardized values.[26] Trust and cooperation "within the data ecosystem is critical to the sustainability of data sharing and relates to issues such as the quality of the data, the level to which the data will be secured after sharing and the responsible use of data by the recipient".[27]

A wide range of data sharing approaches exist in the organizational setting, including open access (allowing unmediated access to publicly shared data), limited access (allowing data access only to those actors who meet particular criteria), and bilateral data sharing (data is only shared between partners). Humanitarian organizations use a variety of methods, platforms and tools in sharing data about affected households and individuals, including email attachments, organizational data platforms like PRIMES, and multiparty computation.[28]

These approaches and methods reflect the diversity of actors involved in cash programming and their respective capacities, resources and programming missions. For those organizations with bilateral or multilateral data sharing agreements in place, interoperability plays a significant role in ensuring that different humanitarian organizations can collaborate effectively, share data and resources, and provide aid to people in need efficiently.[29]

Interoperability ensures that the sharing of information between organizations such as the UNHCR and WFP "is done in a timely, automated, and secure manner bringing better service to" beneficiaries.[30] Given this project's focus on the technical and semantic layers of interoperability, effective technical approaches capable of bolstering the improvement of data sharing among organizations and other stakeholders are thus necessary. For example, the landscaping in Phase 1 identified how the piloting of zero-knowledge proof – a method by which one party (such as CVA recipient) can prove to another party (such as aid or government agency) that a given personal identifying information is true while avoiding conveying additional information beyond the needed facts or details – are important to not only deduplication efforts but also in data security and privacy, especially when encryption is applied.

23 OCHA. 2021 (Ibid.), p. 3.

24 OCHA, UNHCR, UNICEF and World Food Programme. 2018. UN Common Cash Statement (UNCCS): Questions & Answers. https://www.unhcr.org/sites/default/files/legacy-pdf/5f5223f44.pdf

25 UNHCR and World Food Programme. Data systems interoperability and data sharing. Joint Programme Excellence and Targeting Hub. https://wfp-unhcr-hub.org/data-systems-interoperability-and-data-sharing

26 UNHCR. 2019. Minimum Core Assistance Delivery Dataset for Affected Populations. www.unhcr.org/media/minimum-core-assistance-delivery-dataset-affected-populations

27 OCHA Centre for Humanitarian Data. 2020. Responsible Approaches to Data Sharing. OCHA. https://centre.humdata.org/guidance-note-responsible-approaches-to-data-sharing, p.2.

28 OCHA Humanitarian Data Exchange. 2020. https://data.humdata.org/dataset/2048a947-5714-4220-905b-e662cbcd14c8/resource/9db86627-8d8b-435f-8455-6c02802f8fee/download/guidance-note-8.pdf

29 OCHA. 2015. Interoperability: Humanitarian Action in a Shared Space. https://www.unocha.org/publications/report/world/interoperability-humanitarian-action-shared-space#; UNHCR. 2018. Addendum on Data Sharing to the January 2011 Memorandum of Understanding between the Office of the United Nations High Commissioner for Refugees (UNHCR) and the World Food Programme (WFP). Refworld. www.refworld.org/docid/5bbcac014.html; Information Commissioner's Office. 2022. Data Sharing Code of Practice. https://ico.org.uk/media/for-organisations/uk-gdpr-guidance-and-resources/data-sharing/data-sharing-a-code-of-practice-1-0.pdf

30 UNHCR. 2018. Addendum on Data Sharing to the January 2011 Memorandum of Understanding between the Office of the United Nations High Commissioner for Refugees (UNHCR) and the World Food Programme (WFP). Refworld. www.refworld.org/docid/5bbcac014.html

Interoperability framework or 'stack'

Interoperability is a multi-dimensional and socio-technical phenomenon, characterized by legal, organizational, semantic and technical elements that are largely interdependent. Ensuring interoperability is a continuous process that requires formal agreements among actors, organizational structures and roles, effective management of interoperability agreements, change management processes, and practical plans around data quality and control, among others.[31]

- ***Legal interoperability*** ensures that organizations and actors that operate under different legal regimes and regulations are able to work together. This often requires organizations to have clear understandings and agreements about how to navigate differences in legal and regulatory frameworks that bind each actor. This layer includes aspects such as the legal bases for collecting and processing personal data, data protection legal frameworks that can help to identify interoperability barriers both at the local and international level, and provisions around technical standards and protocols.
- ***Organizational interoperability*** refers to the way organizations coordinate their operational processes, responsibilities and objectives to meet commonly agreed and mutually beneficial goals. This involves identifying and integrating various organizational processes and relevant data sharing practices. This element of interoperability also focuses on users and beneficiaries by ensuring user-centric, effective and quality service delivery.
- ***Semantic interoperability*** ensures that the precise format and meaning of shared data are preserved and understood between parties, throughout the data sharing process. This element of interoperability encompasses both semantic and syntactic aspects. The semantic aspect involves the development of vocabularies, code lists, and models to describe data elements and data sharing processes to ensure a common understanding between or among parties. The syntactic aspect involves the description of the exact format of data that's being shared by parties in terms of grammar and format.
- ***Technical interoperability*** involves various applications and infrastructures linking organizational information systems and services. It includes device hardware, data integration services, data formats, communication protocols and interconnection services, among others.

Although the focus of this analysis is on the bottom of the stack, i.e. on the technical and semantic layers, we believe that alignment on organizational and legal issues is actually more difficult to achieve in practice than the technical considerations.

In Figure 1, we present a visual of the different layers of the interoperability 'stack', as well as three cross-cutting themes that must be addressed at each layer: governance, ecosystem and user protection. The horizontal layers of the stack focus on the tactical elements that must be in place for interoperable practices, while the three vertical elements highlight critical considerations that must be addressed for models to be sustainable.

31 European Commission. 2017. New European interoperability framework: Promoting seamless services and data flows for European public administrations. Publications Office. https://data.europa.eu/doi/10.2799/78681

Figure 1. Interoperability 'stack'

System interoperability	Organizational	Ecosystem	User protection
Legal Data sharing agreements and other contracts; regulations at local, national and regional level, including GDPR	To what extent can data standardization lead to more standardized and efficient legal regimes (e.g. DSAs)?	Could different operating models for interoperability create new fiduciary or stewardship roles?	How can privacy policies, access to data, redress mechanisms, and other protections be maintained?
Governance Operations and processes, including workflows, data management, informed consent, user access, security policies	Is it possible to standardize workflows/ SOPs or decision trees around referrals or deduplication?	To what extent can this initiative engage technology vendors to support key system functionality?	How can informed consent and use authorizations be standardized and persist with data as it is shared?
Semantic Definition of data fields and possible values (e.g., gender), relationship/ hierarchy of data objects (e.g. household <> individual)	As likely requires more flexibility than technical standards, will semantic standards sit at country level? Regional?	How might these related to existing SDO (standards development organization) efforts in other sectors?	How can the design consider user input and support localization (e.g. alternative definitions of household)?
Technical Device hardware, communication protocols, data formats (e.g. XML, JSON, HXL), storage, APIs	What is the right entity to drive alignment of technical standards? A new working group?	Assuming an open-source stack, how is a community built to design and support ongoing development of technical components?	How will user needs be prioritized to ensure that the technology used does not exclude already marginalized individuals?

Adapted from European Commission, "New European Interoperability Framework"

Drivers and barriers

The following are some of the factors driving the increasing efforts to achieve interoperability in the humanitarian sector:

- ***The growing digitalization and digital tool adoption among many humanitarian organizations.*** As organizations in the humanitarian sector are realizing the crucial role that digital technologies can play in addressing complex operational challenges, especially around the provision of humanitarian assistance, the need to systematize and effectively integrate different components of their often siloed technology systems is becoming increasingly important.
- ***The adoption of cash as an effective and sustainable modality for humanitarian aid*** requires different technical approaches and operating frameworks compared to in-kind assistance. Given that cash coordination is prone to problems like duplication and diversion, more and more organizations are realizing the importance of data sharing with stakeholders and system interactions with other trusted organizations in the sector to effectively monitor cash flow.
- ***The need for improved engagement and effective collaboration among various private and state actors in the humanitarian field,*** especially in the frontier of information sharing and proactive response. Gray & Colling (2021) highlighted ways the promise of improved engagement of actors and stakeholders (like volunteers, recipients and employees) spurs efforts to achieve interoperability within organizations and across sectors.[32] In addition, as some governments and humanitarian organizations explore the potential for integrating CVA with social protection programming, their need for effective cooperation requires interoperable systems.

32 Gray, B. & Colling, M. 2021. Understanding Interoperability in Humanitarian Aid Organizations. WiP Paper – Enhancing Resilient Response in Inter-organizational Contexts Proceedings of the 18th ISCRAM Conference – Blacksburg, VA, US, May 2021. http://idl.iscram.org/files/brionygray/2021/2345_BrionyGray+MatthewColling2021.pdf

- ***The need to improve the efficiency, effectiveness and quality of humanitarian cash programming.*** Leading organizations in the humanitarian sector recognize the need to improve coordination at country or regional levels, e.g. at the CWG, as well as the need to have systems that are capable of accessing and processing data from multiple internal (and sometimes external) sources to support accurate beneficiary registrations, safeguard individual referrals, and streamline payment processing via financial service providers.[33]

- ***The desire to maintain flexibility and adaptability, and mitigate the problems of depending on one technology and technology supplier.*** Ensuring interoperability as a critical and necessary standard for any technology system within the humanitarian sector is seen as one way of maintaining flexibility in the system and suppliers. The problem of 'vendor lock-in' is increasingly recognized as a barrier to achieving efficiency and effectiveness through technological systems. Similarly, the dominance of a few systems in the humanitarian sector is seen by many as a problem to be solved.

Despite the benefits that could be derived from making systems interoperable in the coordination and delivery of humanitarian cash assistance, several barriers still exist:

- ***Political economy of the sector, including power dynamics and competition for funding.*** The power asymmetries in the sector reflect stark differences in resources and technical capacity, differences that manifest in larger organizations' ability to use more sophisticated technical systems and data management practices. Especially when it comes to securing funding for cash programming, these capabilities are a key competitive differentiator, which creates complicated incentives regarding data sharing and investments in technical interoperability.

- ***Project funding and implementation conditions limit flexibility in system use and data management.*** Humanitarian donors are often unwilling to support technical undertakings of humanitarian organizations that lack direct impact on targeted populations. In addition, some donors specify in their contracts the kinds of data that should be collected and the form it should take, leading to reliance on systems that support these conditions.

- ***Lack of harmonized approaches among actors particularly around beneficiary registration and targeting.*** Data sharing and system integration require coordinated standards and processes, however, many humanitarian field officers are often overloaded with daily operations and constantly changing emergency situations and are therefore left with little to no discretionary time to introduce or implement standards and tools necessary for greater interoperability.[34]

- ***Varied technical capacities across organizations in the sector.*** Efforts to realize technical interoperability across organizations heavily depend on the technical capabilities of actors in the cash delivery ecosystem. As only a handful of organizations (especially large organizations) can invest in sophisticated technologies to facilitate bilateral or multilateral data sharing and system interactions, many other organizations lag behind, which then derails efforts toward effective deduplication and secured individual referrals.[33]

33 Development Initiatives and Publish What You Fund. 2017. The frontiers of data interoperability for sustainable development. www.publishwhatyoufund.org/app/uploads/2017/11/JUDS_Report_Web_061117.pdf; IFRC. 2023; USAID. 2022. Digital Strategy 2020-2024. www.usaid.gov/sites/default/files/2022-05/USAID_Digital_Strategy.pdf.pdf

Current approaches

Salvadoran Red Cross supports families from communities with the humanitarian cash transfer assistance program amidst COVID19. (Photo Credit: Cruz Roja Salvadoreña)

In this section we describe how organizations we spoke with currently approach data sharing for deduplication and referrals.[36] We start with an overview of each of the two use cases, with a focus on the technical architectures and roles and responsibilities of participating organizations. We then present a stylized 'workflow' of how data moves through and across organizations, including references to the typical technical systems and processes used at each stage of the workflow. This workflow model decomposes the processes and technologies used to accomplish data sharing, providing a more nuanced view of where specific barriers and opportunities lie for improved interoperability.

Referrals

When an organization cannot provide the services that an individual or household requires, it will typically refer them to another organization that can. If there is no data-sharing mechanism between those organizations, the beneficiary will typically have to be registered again by the second organization, a duplicative process that is often onerous for the individuals seeking assistance, while also consuming additional time and resources to intake and record information a second or even third time.

To avoid this duplication in the referrals process, the source organization often sends the information that it recorded about the affected person(s) to the recipient organization. Typically, the organization with the source data will create a file that contains biodata (e.g. name, date of birth) and programming information (e.g. assessments or services needed). It will then email it to the recipient organization,

34 UN Statistics Division. 2018. Introduction: Data Interoperability Guide. UN Statistics Wiki. https://unstats.un.org/wiki/display/InteropGuide/Introduction

35 Gray, B. & Colling, M. 2021. Understanding Interoperability in Humanitarian Aid Organizations. WiP Paper – Enhancing Resilient Response in Inter-organizational Contexts Proceedings of the 18th ISCRAM Conference – Blacksburg, VA, US, May 2021. http://idl.iscram.org/files/brionygray/2021/2345_BrionyGray+MatthewColling2021.pdf, p. 444; IFRC. 2023; OCHA Humanitarian Data Exchange. 2020.

36 For a detailed analysis of these use cases, see the Landscape Mapping Reports from Phase 1, https://interoperability.ifrc.org/services-1

which processes the data and updates its own information systems accordingly. Because referrals data is typically low volume and low frequency, and doesn't require automated data processing, the reconciliation function by the recipient organization is typically performed manually.

Of course, bilateral data sharing isn't always via email. It also happens via phone calls, in-person meetings, pen and paper, by exchanging physical storage devices such as a USB drive, or digitally but with colocation of personnel and systems (e.g. staff members from different organizations sitting at a table and comparing spreadsheets). We do not explore these cases in this analysis.

A key functional requirement of the referrals use case is that the actual 'raw' data about the individual or household is usually shared. Compare this with the deduplication use case (below), where in most instances the actual raw data can be shielded or obscured from view by partner organizations.

This is a significant requirement because some referrals include extremely sensitive information about vulnerable individuals, which requires the most strict data protection practices possible. For example, in protection referrals, the case may involve a child or an individual experiencing gender-based violence, and therefore require the strictest security and data protection measures possible.

While a simple bilateral data transfer of some type is the minimum requirement for a referral, many contexts would benefit from more sophisticated case management, e.g. the ability for the source organization to maintain visibility into service provisioning by the recipient organization, or the creation of shared case files that improve transparency and accountability.

Deduplication

Deduplication is the process of comparing lists of intended beneficiaries in order to eliminate any unintentional duplicates. In contrast to referrals, the deduplication use case almost always involves multilateral coordination between three or more organizations – in large responses, such as Ukraine, dozens of organizations may be involved – and thus requires a different operating model to function, as the data must be reconciled across all participating organizations. This typically requires a single registry or database that aggregates the lists from all participants using business logic to determine what counts as a duplicate.

The traditional model for deduplication is a single registry managed by a 'host' organization, which takes on the extra responsibility of hosting the registry or database used to perform the reconciliation process. In most instances the host organization is a member or chair of the CWG and is actively involved in coordinating programming more broadly. But other actors can play the role of host organization; in one case we were told about, the operating organizations send their lists to the regional mobile operator, who then performs the deduplication before sending out payments to mobile money wallets. An example of the single registry, host organization model is UNHCR's PRIMES/ Refugee Assistance Information System (RAIS).

A variation of this approach is a jointly managed registry, where instead of a host organization having admin rights, each participating organization would have equal visibility and access to the collective data, enabling a more equitable governance structure and reducing concerns over lack of transparency. This is what the Building Blocks and GeniusChain solutions are seeking to accomplish through the use of a distributed ledger (i.e. a blockchain), which creates a shared infrastructure where participating organizations can host their own instance of what is essentially a shared registry.[37] Other approaches, like the data stewardship model currently being explored by the CCD consortium, focus more on the governance framework, and not the code itself, to create a system with equal access and responsibility for all participants.

37 While possible in principle, we found very few organizations are actually hosting a node in any distributed ledger scheme. The technical and operational requirements for hosting a node can vary significantly, but could pose a barrier to many NGOs. As a reference, the requirements for hosting a node of the Sovrin network, a distributed ledger used for decentralized identity, include hardware specifications, cyber security protocols, minimum uptime requirements and reporting obligations. See Sovrin. Steward Technical and Organizational Policies V2. https://sovrin.org/wp-content/uploads/Steward-Technical-and-Organizational-Policies-V2.pdf

The actual technical approach used for deduplication can vary significantly. The most sophisticated systems tend to use online environments where the administrators from participating organizations can log in and then upload their files directly to the system, which will then perform processing functions according to the business logic that has been programmed. For example, for deduplication, the system may check the input file against all other existing files and where a duplicate is certain, exclude that individual, where a duplicate is suspected, flag it for manual review, etc. In more rudimentary models, each organization may send a spreadsheet via email to the host organization, which will compile all the lists and perform a deduplication function manually.

Given the participation of so many organizations, and the requirement for automated processing functions that execute agreed-on business rules, the selection process for a technical system to handle deduplication can be complex. For example, the CWG in Ukraine evaluated at least four different solutions.[38]

Typically coordinated by the CWG, deduplication requires all participants to align on semantic definitions (e.g. date of birth versus age) and standard operating procedures (e.g. at what point in the workflow process to submit beneficiary lists, how to handle duplicates) in order to design an effective process for that operating environment.

As a general rule, participants in a single-registry model can only see the data that they contribute, avoiding the need to have data sharing agreements with all the other participants. If the host organization has visibility or access to the aggregate data, participating organizations will usually have a contractual agreement with the host, who will also bear increased responsibilities for ensuring security and data protection.

Different levels of data access and organizational policies toward data protection can result in significant friction in establishing multilateral data-sharing processes. For example, UN agencies typically operate under a distinct regulatory regime – including some legal immunities – compared to international NGOs operating under GDPR, or national NGOs operating under local data protection policies. For organizations under more strict regulation, there can be a perception of increased legal liability for sharing data with an organization that has lower levels of accountability around data protection.

Participation deduplication (and other data sharing use cases) can also be problematic if the approach uses a hosted registry that operates as a 'hub and spokes' model, where multiple (typically smaller) organizations feed data into a singular database which can only be fully seen and accessed by the host. The disparity in access, and lack of control over the systems and outputs, can make the smaller 'spoke' organizations wary of contributing their data. The key concern voiced in this context is that contributing to a central registry only increases the value of that registry (e.g. in the eyes of funders looking for highly scaled operating partners), and thus reinforces the power asymmetry between the host and participating organizations.

Data sharing workflows

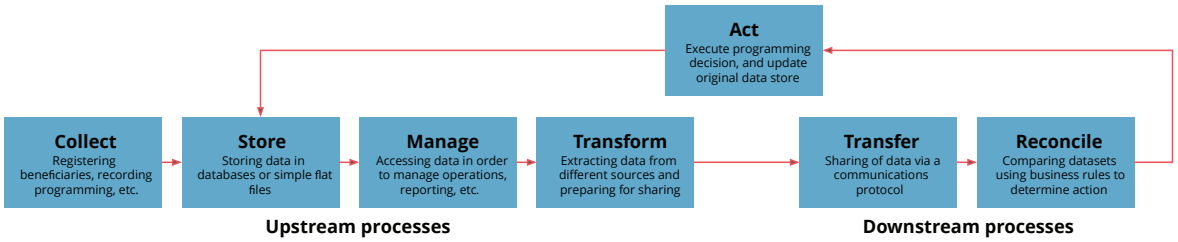
In this section we present a simple data-sharing 'workflow' of how information moves through an organization and is then shared with other organizations. Breaking down the typical workflow helps us to better identify the key activities involved, where there are challenges, and where there might be opportunities for improvements.

The workflow frame creates a distinction between those activities that take place within the organization, here designated as 'upstream,' versus those that take place outside the organization, designated as 'downstream.' While the focus of this analysis is on the downstream activities – the processing and management of data after it leaves the organization – it is important to recognize the role of upstream activities in both the use cases.

38 sUkraine Cash Working Group. 2022. Task Team 3: Deduplication and Registration: Potential Solutions for Deduplication. April 2022. https://reliefweb.int/report/ukraine/ukraine-cash-working-group-task-team-3-deduplication-and-registration-potential-solutions-deduplication-april-2022

For example, with deduplication, many organizations perform some level of deduplication internally as a standard operating procedure, before sharing their data with an external party. Therefore in any holistic evaluation of the best ways to improve deduplication, these upstream activities become salient considerations, if perhaps more challenging to address, as any changes could impact long-established systems and practices.

Figure 2. Stylized data-sharing workflow



Upstream

The workflow activities internal to the organization, here called 'upstream' activities, include typical data collection, storage and management processes. Humanitarian organizations use a variety of technical systems to perform these functions, ranging from basic business software (spreadsheets) to purpose-built beneficiary management tools (essentially, customer relationship management systems (CRMs)) designed for humanitarian response.

Collect

Data collection revolves around registering beneficiaries and recording programming activities, e.g. services provided or assessments. Registration activity typically happens in 'the field' where frontline workers are engaging directly with affected persons, whether at a border crossing, aid distribution point, or other location where individuals seek assistance. Many organizations perform registration activities using smartphones and mobile apps (e.g. KoboToolbox and ODK) designed specifically for this purpose, including offline functionality. In contexts where mobile phone penetration and network reliability are high, such as Ukraine, self-registration is also possible.[39]

One way organizations try to reduce duplicates is through explicit coordination of the registration process. For example, a common practice is to divide registration activities by geography, so that only one organization is working in a given place at a given time. However, duplicates can still occur, especially in displacement contexts or where there are different registration channels (e.g. self-registration via an app as well as in-person registration). In most cases, organizations performing registration will conduct some kind of internal deduplication process to minimize errors and resource costs down the road.

Store

The requirements and challenges of storing humanitarian data are not unique. Some organizations continue to keep their data stored locally on-premises, but many if not most now take advantage of cloud storage to maintain beneficiary and programming records. Organizations that use systems such as RedRose or KoboToolbox will tend to use those vendors' cloud storage options, while some – often those with bespoke systems – use storage offerings directly from the major cloud providers, e.g. Amazon Web Services or Azure. Because registration and other data collection efforts are often conducted offline, many organizations must establish processes to synchronize this cached data with permanent storage.

39 New Zealand Red Cross. 2023. Cash assistance provides vital, dignified support for Ukrainians. OCHA Reliefweb. News and press release, 23 February 2023. https://reliefweb.int/report/ukraine/cash-assistance-provides-vital-dignified-support-ukrainians; IFRC. (n.d). Register with the Polish Red Cross for Cash Assistance. https://ukrainefinancialassistance.ifrc.org/polish-red-cross

One key consideration with storage is the scope of the population included, i.e. to what extent the database reflects a global population of beneficiaries served by the organization versus subsets (e.g. national) of that population. For example, UNHCR's proGres system for beneficiary management was historically decentralized to each local response, with hundreds of unconnected databases across the world. This architecture was changed in 2018 when the organization began a transition to a single, global database of beneficiaries. On this issue, organizations must balance the benefits of unicity against the increased risks of singular data stores.

Manage

Managing beneficiary and programming data involves any number of distinct use cases, from compiling vulnerability assessments to creating donor reports to everything in between. Most of the time, data management occurs within the same system used to store the data. For example, end-to-end beneficiary management systems such as Last Mile Mobile Solutions (LMMS) and RedRose, which are essentially CRM tools, are used for handling virtually all aspects of data management in a typical workflow. On the other hand, smaller organizations that rely on their data collection tool – e.g. Kobo – for storage will often export data to Excel to clean, visualize or otherwise manage the data. Unsurprisingly, having to manually export data to perform data management functions is time-consuming and more prone to errors and security risks.

Transform

The process of extracting and transforming data in preparation for sharing outside the organization will look different for every organization, and depends highly on not only the system(s) used for data management, but also the use case and intended recipient of the data.

For deduplication, the data from each participating organization must be harmonized for the process to work well. In most cases, the CWG or host organization will specify a template (usually a csv file with defined headers) for submitting data; this template will indicate the required fields so that all participants share consistently formatted records. Perhaps because the requirement for consistent data is much less strict with referrals, they are less likely to conform to a standard template used by all organizations, though larger and more established responses may have them.[40]

The transform step plays a critical role in the overall data-sharing process, as it is typically when any agreed data standards or a data schema are enforced. These are typically semantic standards, such as the required data fields (e.g. first name, last name, date of birth), the possible values for those fields (e.g. DD/MM/YYYY), and data object relationships (e.g. max number of members of a household). An unrelated use case for transforming data for external sharing is aggregate reporting of organization activities, for example 4Ws at the cluster level.[41] We explore the importance of data standardization in more detail in the 'Recommendations' section.

Downstream

Downstream activities are those that typically take place outside the organization, i.e. the processing activities that occur once the data has left the organization's systems. As these activities comprise the core functions for each use case, they are the focus of this analysis.

Transfer

The default, and most common, method for transferring data is still via email. Perhaps the next most common method is simple file transfers using FTP or HTTP protocols (e.g. when uploading a file to a web portal). In the context of system-to-system interoperability, the assumption is that APIs are the de facto mechanism for data transfer, but this requires the systems in question to have public APIs, and the interested organizations to have the technical capacity and operational investment of configuring their systems appropriately.

40 See Appendix A for an example from the Ukraine response.

41 4Ws is shorthand for Who does What, Where and When; it is designed as simple yet critical high-level aggregate information for coordinating response efforts.

Reconcile

The reconciliation step encompasses a wide range of different potential processes and activities, from automated deduplication to manual copy and paste of beneficiary information for a referral. All of the single-registry solutions used for deduplication that we analysed support sophisticated functionality to execute business logic according to the use case and programme needs. Designing and developing these functions requires significant coordination and alignment across all participating organizations to agree to semantic standards (e.g. how to record date of birth, or gender) as well as harmonize programming elements (e.g. value of cash assistance per month, number of months for each programme, eligibility criteria). Because there are always edge cases and exceptions in every programme, designing automated functions for reconciliation can become quite complex in contexts with a wide diversity of programmes.

Act

Depending on the use case and what level of automation happens during reconciliation, organizations will need to take action to fully complete the process. For example, in a deduplication exercise the registry host will typically provide a report or output of those individuals or households that require a manual review, and those that are approved to receive the programming service. Part of the coordination and alignment process is to also agree on the standard operating procedures for how to handle different kinds of exceptions or edge cases. In most cases, after reconciliation the source organization will need to update its own records to reflect the service provisioning, referral or other outcome.

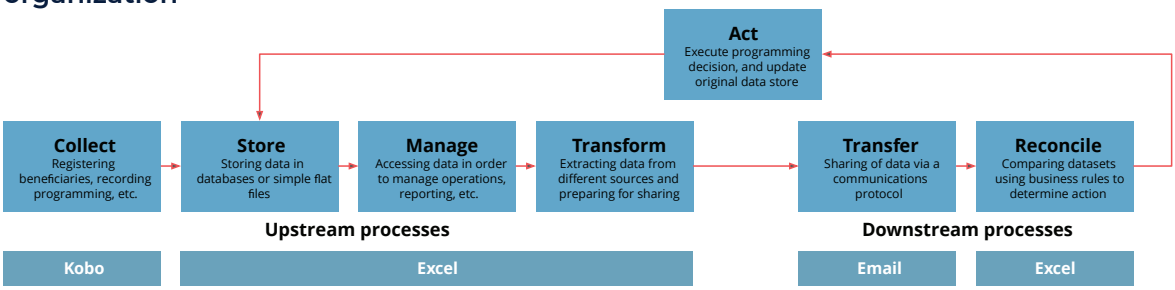
A refugee withdraws financial assistance, which she receives as part of a cash assistance programme in Turkey, from an ATM machine. (Photo Credit: Turkish Red Crescent)

A diversity of systems

In any given context, organizations typically use a combination of systems and processes to perform data sharing, and those systems may vary depending on whether the use case is deduplication, referrals or something else. This diversity makes it extremely difficult to make generalizations about specific approaches to data sharing, as even the same systems can be configured in different ways with different impacts on, for example, data security. But in an effort to frame this wide range of systems and usage profiles, we here describe in broad terms contrasting ends of the spectrum as stylized archetypes to help illustrate this diversity.

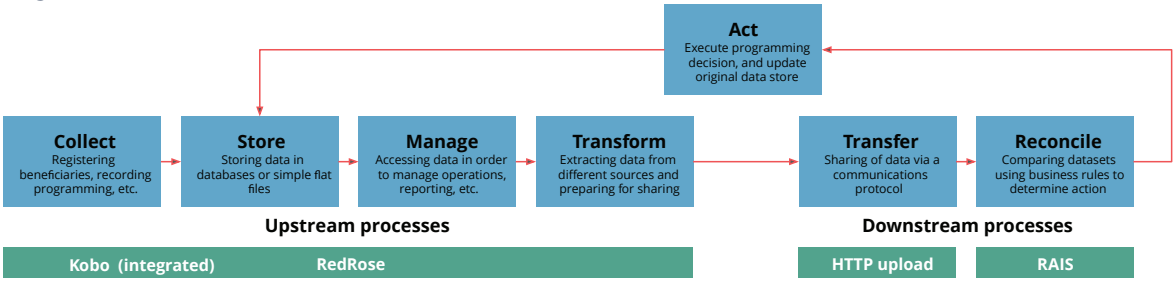
Smaller, lower-resourced organizations are often working exclusively in Excel, or sometimes with a Kobo or ODK mobile application for data collection, plus Excel for data management. To avoid cloud storage costs and complexities, data will be exported from the data collection tool (or typed in from pen and paper collection) and stored as csv or xls files locally, where it is cleaned and processed using Excel. When it comes to time for sharing, e.g. for a referral, the organization creates a csv or xls file with the appropriate information, and sends as an attachment via email to a partner organization, which manually processes or reconciles the data by copying information into its own records.

Figure 3. Stylized example of systems and workflow for a lower-resourced organization



Larger organizations, or those who enjoy the support of an umbrella entity like the International Federation of Red Cross and Red Crescent Societies (IFRC), are more likely to have access to more sophisticated tools, including end-to-end beneficiary management systems like RedRose. These systems typically offer customizable cloud storage options, as well as the ability to integrate with data collection and other tools. In this stylized example, the source organization collects registration data via Kobo, which is integrated into RedRose; the organization extracts the relevant data from its RedRose database, and uploads it as a csv file into UNHCR's online RAIS platform for deduplication.

Figure 4. Stylized example of systems and workflow for a higher-resourced organization



Alternative approaches

2,000 households affected by earthquake in the Tanahun district in Nepal received unconditional cash grants from the Nepal Red Cross, Danish Red Cross. Photo Credit. Danish Red Cross)

In this section we describe how alternative technical approaches, including some emerging technologies, may offer benefits that could improve data sharing in cash programming, in particular around data privacy and protection.

The key distinction for many of these alternative approaches is that they obscure raw data from the recipient organization(s), fundamentally improving the security and privacy of beneficiary data. Not only does this improve data protection, but in some cases it can mitigate the need for data-sharing agreements or other legal contracts, as the recipient organization is not accessing raw data.[42] This can dramatically expedite the process of launching new programmes in the early stages of an emergency, when response time can be critical.

Advanced cryptography

Emerging approaches to cryptography offer new possibilities for sharing information without revealing the underlying raw data. As with most emergent technology, there remains a gulf between what is theoretically possible with these tools, and what has been proven at scale in production environments with actual users.

One promising area of work is around the implementation of homomorphic encryption. This scheme supports the arbitrary manipulation of encrypted data as-is. Without having to procure a decryption key to perform computations it becomes feasible to protect data in-use (in addition to in-transit and at-rest) and eliminate the deficiencies of data between parties. Among the many interesting use cases is the ability to share sensitive data with an organization that will run analytics on top of the encrypted data without revealing their contents.

42 The definition of 'personal information' varies by jurisdiction, making generalizations difficult, but under GDPR, for example, a hashed identifier is still considered personally identifiable information.

Zero-knowledge proofs (ZKPs) use cryptography to enable the validation of information (the 'proof') without revealing the underlying data (without 'knowledge'). The classic example given of a ZKP is to check whether an individual is eligible based on minimum age – with a ZKP approach, the verifying organization would only receive information indicating 'eligible'/'not eligible' as opposed to the actual birth date or age of the individual. In the context of the use cases here, ZKPs could be used in a deduplication process that would enable the organization to validate that an individual is a match/duplicate, without having to actually share information about the individual.

Clean rooms

Some of these emerging technologies are being incorporated into enterprise products for privacy-preserving data sharing. 'Clean rooms' offer customers the ability to invite their partners or other collaborators to jointly contribute each organization's data to a cloud environment, where partners can run queries and data processing functions on each other's data without ever seeing the raw data or having access to it. Companies like Amazon and Habu offer general purpose clean rooms for enterprise customers.

Many clean room solutions require the data from each participating organization to be stored in a centralized environment. But decentralized versions of the clean room model, for example that by Tune Insight, allow each participating organization to maintain full ownership of its data, while the platform coordinates and delegates the data processing across each node such that data never leaves its original location.

User wallets

Distributed ledger technologies have ushered in a new category of solutions that fundamentally change the operating model for how and where data is managed.[43] With traditional models, in the humanitarian sector and elsewhere, each service provider records and maintains data on individuals in their own system. Because each service provider maintains its own records, the result is multiple siloed databases that each have different views of the individual, and require data sharing or interoperability arrangements to jointly process data.

User wallet architectures change this dynamic by storing critical information about the individual in a digital 'wallet' that the individual controls access to. In this way, an organization would ask permission to record information about the individual (e.g. a vulnerability assessment, or services provided) to the wallet, which the individual could then choose to share with other organizations. The data doesn't have to move from one organization to the next, and the individual maintains control – in theory – over who is accessing her data.

43 The DIGID consortium ran pilots using decentralized user wallets for identification purposes in Kenya and Uganda; more details at https://hiplatform.org/digid

Lessons from other industries

Salvadoran Red Cross supports families from communities with the humanitarian cash transfer assistance program amidst COVID19. (Photo Credit: Cruz Roja Salvadoreña)

The humanitarian sector is not unique in its need for improved interoperability in the digital era. Other industries, including financial services, telecommunications and healthcare, have invested heavily in developing interoperable systems for improved service delivery and value to end users.

The healthcare sector is perhaps a useful case to study and draw lessons from, as there are some clear parallels with the humanitarian sector. First, healthcare data about patients is extremely sensitive, and so must be secured and protected to the greatest extent possible. Second, the operating model and use cases in the health ecosystem can mimic a humanitarian response; for example, healthcare organizations often refer patients to other healthcare organizations, and must maintain case files for continuity of care. And third, the healthcare industry is composed of a range of nonprofit and private-sector actors that must share data, representing different levels of resourcing and capacity, and operating across different legal jurisdictions and data protection regulations.

To complement the main technical analysis, we therefore also studied how standardization in the healthcare sector – specifically, the Fast Healthcare Interoperability Resources (FHIR) standard – has been successful in driving interoperability, and any lessons learned for the humanitarian space.

About HL7 and FHIR

The FHIR standard is developed and maintained by HL7 (Health Level Seven International), a nonprofit organization that acts as a formal standards-setting organization for healthcare data exchange.[44] In the FHIR framework, data sharing can occur between various entities such as healthcare providers, payers, patients and other stakeholders involved in the healthcare ecosystem.

44 HL7. FHIR. Fast Healthcare Interoperability Resources. www.hl7.org/fhir/overview.html

The FHIR standard is a sophisticated, comprehensive specification for data exchange. The core consists of data objects called 'resources,' which include definitions (e.g. description of the content), datatypes (e.g. date, Boolean), metadata (e.g. last updated), and a human-readable summary. FHIR-compatible content can be represented in a number of actual formats, including JSON and XML. While FHIR was designed to support REST APIs, other approaches are possible, including a service-oriented architecture model.

Lessons learned

The nonprofit housing the FHIR standard, HL7, has been building and promoting standards for interoperability since its founding in 1987, and standardization in the healthcare industry is significantly more mature than the more nascent efforts in humanitarian aid. Nevertheless, there are a number of key lessons from the HL7 experience and successes that are relevant to the humanitarian sector. We summarize a few highlights according to whether they relate to process and governance, or the technical elements of the standard itself.

Process and governance

- **Choose adoption over accreditation** – The initial standard was explicit in prioritizing usage in production environments before seeking formal accreditation from standards-setting bodies; as it matured, later versions are fully accredited by the International Organization for Standardization (ISO), the American National Standards Institute (ANSI) and other institutions.
- **Engage the private sector** – Much of the early innovation was driven by vendors looking for cost savings as a competitive differentiation when doing custom implementations; HL7 did have to carefully balance those interests to ensure it was perceived as neutral.
- **Build on what exists (without playing favourites)** – The original specification stated "The Standard should be built upon the experience of existing production protocols. It should not, however, favor the proprietary interests of specific companies to the detriment of other users of the Standard".

Technical

- **Define a limited core standard, design for extensibility** – Recognizing that trying to design a standard that covers all use cases and contexts leads to a cumbersome and complex specification, FHIR defines core requirements and then uses 'extensions' that allow implementers to define custom data resources and definitions according to their needs.
- **Enable a range of modalities** – While the FHIR standard specifies the definitions and datatypes of the information exchanged, it supports multiple data formats and transmission mechanisms, enabling different technical architecture and systems.
- **Include human-readable tags** – As a backup measure, the FHIR standard specifies the inclusion of human-readable summaries with each data resource, providing a mechanism for a practitioner to parse a data resource and understand its meaning even if no computer processing is available.

The success of FHIR and HL7 standards in the healthcare industry is just one example of standards-driven interoperability, and there are undoubtedly other examples that would offer relevant insights. A deeper study of the early stages of successful standardization efforts could be helpful context for where the humanitarian sector finds itself today.

45 Ringholm. 2014. The early history of health Level 7. Whitepaper. www.ringholm.com/docs/the_early_history_of_health_level_7_HL7.htm

Evaluation

A woman affected by extensive flooding due to heavy monsoon rains in the Da Teh district in Viet Nam is receiving Multipurpose Cash Assistance from the Viet Nam Red Cross Society. (Photo Credit: IFRC)

In this section we compile information on the use cases, data workflows and system functionality to develop a simple framework for evaluating different technical approaches. Given the wide diversity in technical systems and how they are configured, we start by categorizing those systems into three high-level operating models. The characteristics of each operating model determine to some extent the technical solutions that can be employed, but in practice each operating model can use a wide range of different technical systems and approaches to sharing data.

We then apply the interoperability principles endorsed by the major donors in the DCF as a lens, deriving five key dimensions relevant to this technical assessment (with supporting text drawn from the 'Donor Cash Forum Statement and Guiding Principles on Interoperability of Data Systems in Humanitarian Cash Programming'):[46]

- **Security and data protection** – "[A]ppropriate, strong, and effective safeguards for the privacy and security of data, as well as mechanisms for informing data subjects about how their data will and will not be used, should be built into all interoperability initiatives."
- **Benefits to beneficiaries** – "The needs and rights of affected populations should be prioritized across the data lifecycle… Giving the communities from which data is collected as much ownership as possible should be an overall goal."
- **Technical accessibility** – "The goal of interoperability should be to facilitate the access, participation, and inclusion of a wide range of organizations… [and] should be designed to facilitate entry of new actors over time with minimal burden."

46 DCF. Donor Cash Forum Statement and Guiding Principles on Interoperability of Data Systems in Humanitarian Cash Programming www.calpnetwork.org/wp-content/uploads/ninja-forms/2/DCF-Interoperability-Statement-FINAL.pdf

- **Open architecture** – "Interoperability initiatives should aim to allow for a distributed or federated network… Where possible, use of open standards, open data and platforms that prevent lock-in to a specific proprietary technology will facilitate the integration of a range of different systems and actors over time, especially those with fewer resources or access to technology."
- **Collective governance** – "Cooperation between organizations should include… the governance of data sharing at the local and possibly global levels – where possible and appropriate, this should be through a formal collective forum."

While this evaluation includes high-level considerations for security and data protection, it is not intended to be an in-depth review. Instead, a separate report focused on security risks and vulnerabilities of data-sharing practices will be published in parallel by the DIGID consortium.[47]

Bilateral data sharing

Typically the norm for referrals, the bilateral data-sharing model comprises all approaches where a single source organization shares information with a recipient organization via a digital channel. This usually means the emailing of xls or csv files from one organization to another. Sometimes, the file will be password-protected, or even encrypted, but often not.

Most practitioners with operational experience understand how pervasive this approach is, and how risky: once an unprotected file has left the source organization, there are no controls available to prevent unauthorized access.

Encrypting the file with a strong key before sending provides a clear improvement in security, mitigating attacks or unauthorized access if the email were to be intercepted. But the file is still sitting on email servers, perhaps on local storage, without any visibility into who might have access.

But bilateral data sharing can also occur using more secure technical approaches, such as a shared platform or online environment for uploading data. If that platform has strong encryption and user access controls (i.e. requires user authentication with customizable permissions), that would provide significant security and data protection benefits to the two sharing organizations. It's important to note, however, that commercial service providers may be subject to law enforcement or other government requests, such that privacy rights and data protection considerations can ultimately depend on the jurisdiction of the hosting firm.

Security and data protection: Default approach using email is the least secure method available. Encryption offers significant improvements. A shared platform or hosted environment offers even better security due to user access controls and permissions, though the jurisdiction of the host matters for privacy and data protection considerations.

Beneficiary benefits: None in any configuration.

Technical accessibility: Emailing files is by far the most accessible approach, and available to almost all organizations. Encrypting files in Excel doesn't add significant complexity. Sharing files through a shared platform or web portal adds minor complexity.

Open architecture: Sending files via email is a relatively distributed and open approach, without any centralized system and using standard internet protocols. A web portal or online environment could use any range of open-source code and standards.

Collective governance: Given the bilateral nature of the sharing, activity is typically governed by bilateral, not multilateral, governance and data-sharing agreements. A web portal would require collective alignment on data standards, possibly standard operating procedures, and probably data-sharing agreements/contracts.

47 Please see https://interoperability.ifrc.org for the latest publications.

Shared registry

This model refers to a shared registry or database that serves as the single source of truth for participating organizations. This is the de facto approach for deduplication, which requires multilateral cooperation and sharing of beneficiary lists. However, there are multiple ways in which a shared registry can be structured or configured that have a significant impact on its operations, including data protection and security.

As described earlier, one important consideration is who runs the registry – i.e. whether there is a 'host' organization fulfilling an administrative role, typically by managing the registry infrastructure, coding the business rules functionality, or in some cases, manually performing the reconciliation process itself, or whether the registry is somehow co-owned and operated by participating organizations. The former approach is exemplified by UNHCR's RAIS system, whereas the latter is what Building Blocks, GeniusChain, and the CCD's data steward model aspire to. While this is more of a governance consideration than a technical one, there are data protection implications in instances where the host organization has access to the data.

Another key consideration is whether the technology used for the registry blinds or obscures the data from view by other participating organizations (including the host, if applicable). This approach only serves some use cases, notably deduplication, where participants don't actually need to see the underlying data, they only need to know the result of matching. Some shared registries in use today employ hashing techniques to obscure the raw data (e.g. name, date of birth) and only conduct the reconciliation processes (matching) on the hashes. Other emerging technologies, such as zero-knowledge proofs (ZKPs), homomorphic encryption, and complete solutions like clean rooms, enable similar functionality. This ability to enable processing on shared data without revealing the underlying information offers obvious benefits in terms of security and data protection. It also can mitigate or eliminate the need for data-sharing agreements or other time-consuming contractual processes. However, this approach doesn't work for all data sharing use cases: for example, with referrals, the recipient organization typically needs to have the underlying raw data (e.g. vulnerability assessment, records of services provided) to effectively process and act on the referral.

Different shared registries may employ different transfer mechanisms. In some cases, the host of a registry will simply receive emails with files attached, introducing all of the security and data protection risks highlighted above. More secure registries will only accept files that have been encrypted and uploaded through a secure web application or portal.

A person receives unconditional cash assistance to rebuild their shelter after an earthquake in Tanahun district in Nepal. (Photo Credit: Danish Red Cross)

Such registries will also typically require user authentication, i.e. the user must log in to the portal before access is granted to upload a file. Such user access controls are a fundamental security practice with enterprise software generally, and offer a host of benefits, including the ability to set granular permissions (e.g. what a user can view, edit, export) and revoke access, useful when a user leaves an employer or changes roles. A related benefit of access controls is logging functionality, where the system records an event each time a user logs in, accesses a file, or performs any action on the platform. Such logs are part of basic security practices, and provide increased transparency and auditability for stakeholders such as donors.

Security and data protection: Registries with user access controls and that require direct uploads offer significant benefits over those with more manual sharing (emailing files). Technologies that obscure raw data dramatically increase data protection and security.

Beneficiary benefits: None in most configurations, with the exception being the data steward model proposed by the CCD consortium, which would provide beneficiaries with the ability to view data collected about them.

Technical accessibility: Straightforward for most participating organizations. Hosts may face significant design and configuration challenges.

Open architecture: Building Blocks and GeniusChain both use a decentralized ledger, which in principle offers security benefits over a singular central registry. No open-source solutions in this category were identified.

Collective governance: If the host can see data, it would require data-sharing agreements across the board. Technologies that obscure raw data can reduce/eliminate the need for data-sharing agreements. All registries with business logic require harmonization on data standards and standard operating procedures.

User wallets

A user-controlled wallet model represents a fundamental departure from traditional operating models and technical architectures. Instead of data residing with each service provider, and thus having to be shared across those organizations, the data about the beneficiary resides in a wallet that the beneficiary controls. In this system, the user typically has to grant access to that wallet each time an organization wants to access information, ensuring that the user has control over how and when her personal data is being used.

There are compelling benefits when the single source of truth about an individual is a personal data store that the user controls access to. For one, the individual has, in principle, greater visibility, agency and control over access to her data. Another key benefit to the user is that a wallet offers portability of the data; i.e. the individual can take her records of services provided or credentials with her after the relationship has ended. And there are clear data protection and security benefits of not having to transfer data from organization to organization – it all stays in one place.

However, this individual agency comes with costs. Any model that gives the individual authority to make decisions about access to her data is also placing the responsibility for such decisions on the individual instead of the organization. Evaluating and managing automated requests for one's data is an unreasonable burden even for highly digitally literate individuals not in a vulnerable state; for persons in humanitarian crisis, this is fully unrealistic.

There are also operational constraints to a wallet model. In contrast to the different variations of a shared registry, or bilateral data sharing, adopting a wallet model requires an organization to change not only downstream processes but also upstream processes. In effect, it requires new practices at multiple stages of the data workflow, from data collection/registration (when a credential would be recorded to the wallet) to service provisioning (when a record of provided services would be recorded to the wallet) to the actual data sharing (when the source organizations would need to notify recipient organization(s) to submit a request to the individual to access the records in the wallet).

In terms of use cases, the wallet model appears to be an excellent solution for referrals, offering data privacy and portability of data when a beneficiary seeks out services from another participating provider.[48]

But for other use cases, primarily deduplication, the suitability is less clear. In a typical deduplication response, multiple organizations compare lists of potential beneficiaries and the assistance they are receiving in order to reduce unintended duplicates. In a true wallet model, the programming information about having received assistance would reside in the individual's wallet, and require her permission to access. Gaining permission from all relevant individuals would clearly be impractical at scale and in the timelines needed for timely distribution of benefits. One alternative would be for each organization to share its own internal records of provided assistance with partner organizations, effectively creating a parallel data workflow that processes deduplication in a standard fashion. But this would seem to violate a central premise of the wallet, that it puts the user in control over who is accessing her information at any point in time.

Security and data protection: In principle, maintaining personal data in the wallet offers security benefits, but the technical architecture is still in its infancy.

Beneficiary benefits: Significant benefits in terms of agency, visibility and portability of data, but with added new responsibilities and liabilities that may outweigh those benefits.

Technical accessibility: High barriers to usage by organizations, as it requires revamp of upstream and downstream data processing activities.

Governance: Requires new governance frameworks for enabling beneficiaries to control their wallets, but in theory reduces or eliminates needs for data-sharing agreements across organizations.

48 A summary of the DIGID consortium's pilot of using user wallets for identification in Kenya and Uganda is available at https://hiplatform.org/digid

Recommendations

A representative of a household affected by an earthquake in Tanahun district, in Nepal displays his identifier for receiving unconditional cash assistance. (Photo Credit: Danish Red Cross)

In this section we continue to apply the DCF principles of interoperability, using them as a lens for interpreting the learnings and insights from this phase of the project into recommendations for how to improve data sharing in the humanitarian sector. We present two key high-level insights from the research, how they align with the DCF principles, and our recommendations for how to take action.

One very clear finding from this technical landscaping is the tremendous diversity in systems and processes used to share data across organizations, and the prominent role that rudimentary practices – i.e. emailing Excel files – still occupy in the sector. While most people recognize the security risks of these practices, there are good reasons why spreadsheets are still used, and it's unrealistic to believe that the organizations that rely on Excel today will migrate to using an API tomorrow. To restate one of our fundamental assumptions, there is no singular technical solution that will work for all organizations, contexts and use cases. There is simply too much diversity in the level of resourcing, technical capacity, operational environment and programming needs.

This view aligns with the DCF principle 'design for inclusivity', where "[t]he goal of interoperability should be to facilitate the access, participation, and inclusion of a wide range of organizations," including those with lower levels of resourcing and capacity. We believe that any recommendations to data-sharing improvements need to balance the technical ideal state with the operational realities of organizations serving beneficiaries in crisis settings. Framed this way, we can shift from thinking about a singular ideal application, and instead embrace foundational tools and best practices that can be relevant to a range of different organizations and operating contexts.

The second key finding was the absence of any widely accepted set of data standards, despite strong demand for them. Instead of a general reference standard, every country context we engaged with had defined its own semantic standards for sharing data within that ecosystem. Sometimes this was an

explicit process coordinated by the CWG, in other instances it was more organic. Yet in virtually every consultation, from operating organizations to UN agencies to private-sector technology vendors, the experts we consulted would strongly welcome a data standard to enable more effective data-sharing practices.

Unsurprisingly, one of the DCF principles is 'standardization and harmonization', where "[o]rganizations should cooperate to ensure the standardization of data structures, exchange methods, and a minimum amount of data variables, including common identifiers, to facilitate the safe sharing of data." While the value of data standardization seems to be widely agreed on, the challenge is how different actors might actually cooperate: who gets a voice, how decisions are made, and what scope is defined can all be potentially contentious questions.

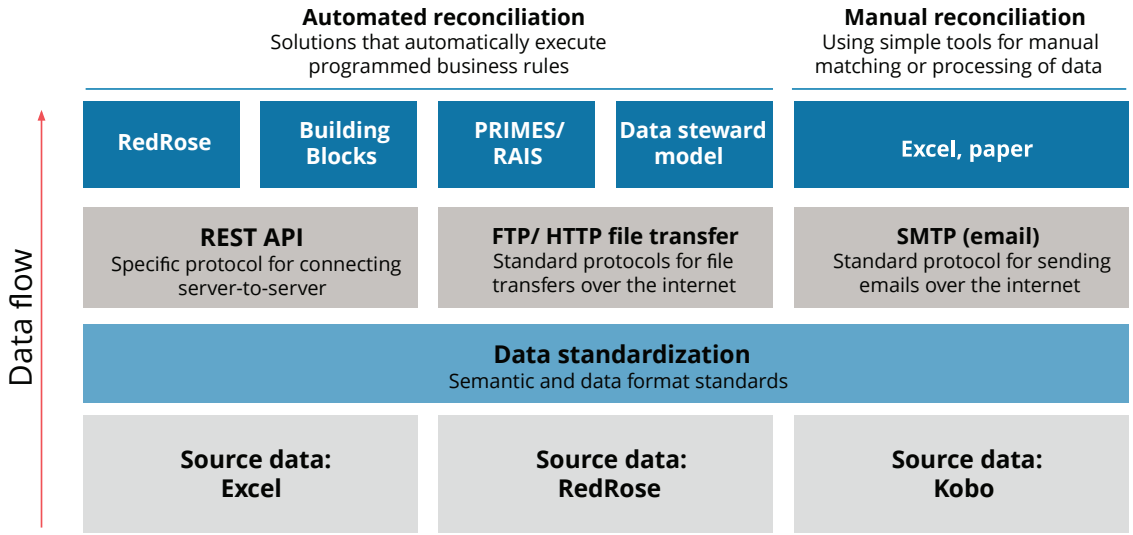
Building on these key findings, we present here two high-level recommendations that we believe are tightly aligned with the DCF principles, and have the best likelihood of advancing the state of data sharing in the sector: advancing a semantic data schema, and developing simple open-source tooling for data sharing.

Advancing a semantic data schema

In any ecosystem, the standardization of data semantics and formats, e.g. a data schema, is the foundation for interoperability. If we accept the premise that there will continue to be a plurality of solutions for collecting, storing, managing and reconciling data, then data standardization is a precursor to enabling the wide variety of systems to each ingest and process other data easily and securely. This applies across the range of different data-sharing practices, including actual system-to-system integrations for data transfer.

To visualize the foundational role that standards play in data sharing, we revisit the workflow model described earlier, but with the axes reversed. This view highlights how the plurality of source data systems, and a plurality of transfer protocols and reconciliation systems, can all benefit from alignment at the data standardization layer. In the earlier workflow model, this is the 'transform' step, when data from internal systems is readied for sharing with outside parties. Creating tools that enforce established data standards at this step would improve downstream data sharing regardless of what systems are being used.

Figure 5. Data sharing workflow, showing foundational role of standardization



Therefore our recommendation is for an investment in the development of a core set of data semantic standards, or data schema, that could serve as the default reference for data-sharing use cases. Of course, given the diversity of programming needs and operational contexts, any standard would need the flexibility to accommodate custom or localized data types and definitions, but our research suggests that a core dataset is possible.

There are precedents in this area. Most notably, the 'Minimum Core Assistance Delivery Dataset for Affected Populations'[49] defines a set of data fields, their descriptions and associated metadata. Arising from the 2018 addendum to the data-sharing agreement between UNHCR and WFP, this 'core dataset' could form the basis for an expanded schema. Another example comes from the Data Sharing working group of the CCD consortium, which has defined its own standard data fields for sharing across CCD members.[50]

Defining the actual fields and descriptions is likely straightforward. Much more difficult will be determining the optimal process for stakeholder engagement and decision-making, identifying the appropriate institutional home for the schema, and developing the right incentive structures to spur adoption. Again, there are precedents in this regard, with the Humanitarian Data Exchange[51] and HXL, the Humanitarian Exchange Language,[52] hosted by UN OCHA. The experiences of the OCHA team in their effort at establishing HXL as a global standard for humanitarian data would also be invaluable inputs to a future standardization effort (note that HXL is intentionally limited to aggregate data types, without any coverage of individual or personal data, so is not applicable to the data sharing use cases explored in this project).

And there are useful references outside the humanitarian sector as well. The Open Identity Exchange (OIX), a nonprofit trust framework provider supported by US government agencies and private-sector firms, recently released its version of data standards for digital identification;[53] while specific only to identity, the schema includes international standards for common biodata fields that could be relevant.

Open-source tooling for data sharing

In larger, well-resourced humanitarian responses, there are typically established solutions for downstream data sharing, including UNHCR's RAIS, WFP's SCOPE or Building Blocks, or a beneficiary management solution such as Last Mile Mobile Solutions (LMMS) or RedRose. As described earlier in this report, these solutions offer general security and access controls that provide significant benefits over rudimentary file sharing practices.

Yet those established solutions are only used for a fraction of total data-sharing instances, typically in 'hub and spoke' models. For example, UNHCR may share beneficiary targeting lists with implementing partners through a formal process using its RAIS system (the 'hub'). But when those partners – especially those with lower levels of technical capacity – need to coordinate with each other (the other 'spokes'), or even within their own organizations, that information won't go through RAIS – it's most likely emailed.

There is therefore a significant gap in terms of systems and tools that lower-resourced organizations can use for data sharing that occurs outside formal structures and systems. Addressing this gap by supporting the development of open-source tooling for data sharing could have a direct impact on improved data-sharing practices.

Our research suggests that one potentially high-value tool would be a simple web portal or cloud environment that would provide an alternative to emailing a spreadsheet. The portal would allow an organization to upload a file and designate sharing permissions with another organization(s),

49 UNHCR. 2019. Minimum Core Assistance Delivery Dataset for Affected Populations. www.unhcr.org/media/minimum-core-assistance-delivery-dataset-affected-populations
50 CCD Data Sharing Agreement. https://drive.google.com/drive/folders/1Q1VeB9UjSz8e3f9xHrs32cmLN4x1CCuA
51 OCHA Humanitarian Data Exchange. https://data.humdata.org
52 Humanitarian Exchange Language (HXL). https://hxlstandard.org
53 OIX Open Identity Exchange. https://openidentityexchange.org/networks/87/item.html?id=700

providing modern security protections without adding significant workflow friction or technical literacy requirements. Furthermore, a portal could easily include functionality to designate and enforce a data schema, providing a simple mechanism for helping organizations adopt and take advantage of the benefits of standardized data.

In considering how this type of tool could be brought to market, we revisit the DCF principles on interoperability, which are clear about the importance of open-source systems for facilitating access and innovation: "Where possible, use of open standards, open data and platforms that prevent lock-in to a specific proprietary technology will facilitate the integration of a range of different systems and actors over time, especially those with fewer resources or access to technology."

Of course, having an open-source code base is not enough to make a solution fit-for-purpose or superior to a proprietary system. But open technologies that are well designed and have robust community engagement – often, including private-sector actors – have great potential for creating more transformative and equitable ecosystems. The 'digital public good/infrastructure' movement, spearheaded by a coalition of organizations including UNICEF, UN Development Programme, multiple bilateral aid organizations, and nonprofits such as the Digital Impact Alliance,[54] Digital Public Goods Alliance[55] and Co-Develop Fund,[56] has done well in highlighting the public interest benefits that open software can offer, especially to organizations and state institutions trying to serve the public good.

And there are clear examples of how transformative open solutions can be in the humanitarian sector. The development of ODK and KoboToolbox as open-source solutions for data collection demonstrated the latent demand for simple-to-use tools that provide modern functionality and security features yet remain accessible to low-resource organizations.

A key characteristic of the Kobo/ODK solution set is that they were designed as modular tools with a limited scope; compared to comprehensive beneficiary management systems, they are focused on a specific use case. This modularity enables organizations to take more flexible approaches in how they construct their backend systems and workflows, adopting specific point solutions when they make sense and not having to commit to a singular monolithic system that may not be the best fit across the board. An example of digital public infrastructure, MOSIP (Modular Open Source Identity Platform) embraces the modularity approach, enabling government agencies flexibility in how they integrate their systems into a modern identification platform.[57]

An open-source, modular tool for downstream data-sharing activities – a 'KoboToolbox for data sharing' – could therefore not only provide direct benefits to many organizations that are currently stuck with rudimentary systems and practices, but could also help to stimulate the ecosystem of service providers and open communities that support such systems.

In this section, we've described how key findings from this research align with tenets of the DCF principles on interoperability, and suggested two recommendations for how those ideas could be put into practice. The establishment of a core data schema for sharing beneficiary data is the critical foundation that can help any further efforts at interoperability be more productive. The development of simple, open-source tools for data sharing could directly serve organizations currently excluded from using modern systems, and could also serve as a mechanism for operationalizing the data schema.

54 Digital Impact Alliance. Digital Public Goods Charter. https://dial.global/work/charter-for-digital-public-goods
55 Digital Public Goods Alliance. Unlocking the potential of open-source technologies for a more equitable world. https://digitalpublicgoods.net
56 Co-Develop. www.codevelop.fund
57 Mosip. https://mosip.io

Conclusions

Oksana, artist from Lviv is receiving a cash card at the Red Cross in Warsaw. (Photo Credit : British Red Cross)

There is widespread agreement on the risks presented by rudimentary – yet still common – methods of data sharing. No one believes that emailing spreadsheets is a good idea.

Yet these challenges are not new. In 2015, an ECHO-funded analysis on the 'Challenges and the State of Play of Interoperability in Cash Transfer Programming' suggested key barriers and potential next steps that remain relevant today.

Progress on this issue is hard, in large part because while the issue is situated within a technical environment, the most significant challenges are not technical in nature. The real barriers to more interoperable systems in humanitarian aid revolve around the political economy of the humanitarian sector itself, the structure of traditional grant funding and procurement mechanisms, and the competition for increasingly commodified – i.e. cash service delivery contracts.

Put another way, there is no technological solution that can solve the significant power imbalances that exist between large UN agencies and smaller NGOs. When robust data (i.e. beneficiary lists) and software systems are seen as competitive differentiators in the allocation of funding, the result is an incentive structure that is diametrically opposed to system interoperability and data sharing.

Real progress on improving interoperability will therefore only be realized if key actors – in particular, donors – can reshape incentives to reward investments in interoperability. This will undoubtedly have to be tied to funding in some way.

58 UNHCR and World Vision. 2015. Improving Cash-Based Interventions Multipurpose Cash Grants and Protection. Enhanced Response Capacity Project 2014–2015. Challenges and the State of Play of Interoperability in Cash Transfer Programming. www.calpnetwork.org/wp-content/uploads/2020/01/erc-executive-summary-interoperability-web.pdf

To help advance our collective understanding of the problem space, we have presented in this report multiple conceptual models of how data sharing for cash programming actually occurs in practice. By mapping how different actors, systems and processes interrelate to produce the practices of sharing data across organizations, we hope to provide a more nuanced understanding of where the challenges and opportunities for improvement can be found.

And finally, rather than advocate for a specific technical approach, this report made the case that there is no singular solution that will work for all organizations in all contexts; capacities, requirements and environments are simply too diverse. Instead, we recommend that stakeholders focus on data standardization as the foundation for improved data sharing and system interoperability, regardless of the technical approach chosen. Put another way, instead of searching for a silver bullet at the application layer, building alignment around open standards at the semantic layer offers the best opportunities for long-term success.

BIBLIOGRAPHY

Allen, D. K., Karanasios, S. & Norman, A. 2014. Information sharing and interoperability: The case of major incident management. *European Journal of Information Systems,* 23(4), 418–432. https://doi.org/10.1057/ejis.2013.8

Brown, I. 2020. *Interoperability as a Tool for Competition Regulation.* Open Forum Academy. https://openforumeurope.org/publications/ofa-research-paper-interoperability-as-a-tool-for-competition-regulation

Currion, P. 2022. *Data Portability and Digital Identity in Humanitarian Aid: A Desk Review.* CCD. www.collaborativecash.org/_files/ugd/477045_e582d8bf08a74b1ab5891b3bc5389bb0.pdf

Data Interoperability Collaborative. 2017. *Improving Data Interoperability for the SDGs.* www.data4sdgs.org/blog/improving-data-interoperability-sdgs

Development Initiatives and Publish What You Fund. 2017. *The frontiers of data interoperability for sustainable development*. www.publishwhatyoufund.org/app/uploads/2017/11/JUDS_Report_Web_061117.pdf

European Commission. 2017. *New European interoperability framework: Promoting seamless services and data flows for European public administrations.* Publications Office. https://data.europa.eu/doi/10.2799/78681

Gal, M. S. & Rubinfeld, D. L. 2019. Data Standardization. *New York University Law Review, 94.*

Gray, B. & Colling, M. 2021. *Understanding Interoperability in Humanitarian Aid Organizations*. WiP Paper – Enhancing Resilient Response in Inter-organizational Contexts Proceedings of the 18th ISCRAM Conference – Blacksburg, VA, US, May 2021. http://idl.iscram.org/files/brionygray/2021/2345_BrionyGray+MatthewColling2021.pdf

Information Commissioner's Office. 2022. *Data Sharing Code of Practice*. https://ico.org.uk/media/for-organisations/uk-gdpr-guidance-and-resources/data-sharing/data-sharing-a-code-of-practice-1-0.pdf

Madon, S. & Schoemaker, E. 2021. Digital identity as a platform for improving refugee management. *Information Systems Journal,* 31(6), 929–953. https://doi.org/10.1111/isj.12353

Mistry, P., Maguire, D., Chikwira L. & Lindsay, T. 2022. *Interoperability is more than technology.* The King's Fund. www.kingsfund.org.uk/publications/digital-interoperability-technology

O'Brien, S. 2017, January 13. *This is how we build a stronger, data-driven humanitarian sector.* World Economic Forum. www.weforum.org/agenda/2017/01/this-is-how-we-build-a-stronger-data-driver-humanitarian%20sector/

OCHA Centre for Humanitarian Data. 2020. *Responsible Approaches to Data Sharing.* OCHA. https://centre.humdata.org/guidance-note-responsible-approaches-to-data-sharing

OCHA Centre for Humanitarian Data. 2021. *OCHA Data Responsibility Guidelines 2021.* https://centre.humdata.org/the-ocha-data-responsibility-guidelines

OECD. 2021. Data Portability, *Interoperability and Digital Platform Competition.* www.oecd.org/daf/competition/data-portability-interoperability-and-digital-platform-competition-2021.pdf

Riley, C. 2020. Unpacking interoperability in competition. *Journal of Cyber Policy,* 5(1), 94–106. https://doi.org/10.1080/23738871.2020.1740754

UN Statistics Division. 2018. *Introduction: Data Interoperability Guide*. UN Statistics Wiki. https://unstats.un.org/wiki/display/InteropGuide/Introduction

UNHCR. 2018. *Addendum on Data Sharing to the January 2011 Memorandum of Understanding between the Office of the United Nations High Commissioner for Refugees (UNHCR) and the World Food Programme (WFP).* Refworld. www.refworld.org/docid/5bbcac014.html

UNHCR. 2023. *World Bank-UNHCR Data Sharing Agreement to Improve Assistance to the Forcibly Displaced.* 27 June 2023, press release. UNHCR UK. www.unhcr.org/uk/news/press-releases/world-bank-unhcr-data-sharing-agreement-improve-assistance-forcibly-displaced

USAID. 2022. *Digital Strategy 2020-2024*. www.usaid.gov/sites/default/files/2022-05/USAID_Digital_Strategy.pdf.pdf

APPENDIX A.
DATA-SHARING TEMPLATE EXAMPLES

Inter-agency referral form, Ukraine

Priority	
Urgent (within 24 hours)	Normal (within 3-5 days)
Case information (if consent has been obtained): *Insert the basic biodata and contact information which is needed for the case to access the required service. Check Service Mapping to see if additional requirements are necessary to access the service.*	
Name (or Case Code if confidential):	Date of birth:
Gender:	Preferred method of contact:
National ID:	Preferred day/time of contact
Address (street/settlement):	
Address 2 (Oblast/Rayon/Territorial community):	
Phone:	Additional (disability etc.):
If a child (under 18 years old) *Remember: if it is not appropriate to involve the child's caregiver (for instance if the caregiver is involved in the abuse), informed assent should be sought from the younger child.*	
Name of primary caregiver:	Relationship to child:
Contact Info for caregiver:	Is child separated or unaccompanied? ◯ Yes ◯ No
Caregiver is informed of referral? ◯ Yes ◯ No (If no, explain):	
Special Note/Specific Need:	
Reasons for Referral: *Describe the minimum information required by the receiving agency to be able to respond to the referral. This can include problem description, whether they receive other assistance, number in the household, etc. For referrals to GBV, CP and Protection case management, do not provide details of the case or incident.*	

Need for service/assistance	
Protection	Health
SGBV	Shelter
Child Protection	Non-Food Items
Legal Assistance	WASH
Multi-Purpose Cash	Nutrition
Education	Other
Food Security and Livelihoods	
Please explain any need for service, and any already provided:	

Consent to Release Information *(read with client/caregiver and answer any questions before s/he signs below)*

I _____ (person of concern name), acknowledge that the service provider, _____ (service provider name) has clearly explained the procedure of the referral to me and has listed the exact information that is to be disclosed. I understand that my information will be treated with confidentiality and respect, and will only be shared as needed to provide assistance and may be used for purposes of humanitarian analysis. By signing this form, I authorize this exchange of information to the specific service provider/s for the specific purpose of providing assistance to my family and/or myself.

Signature (or caregiver if child):	Date:
Referred By:	**Referred To:**
Name:	Name:
Sector of operation:	Sector of operation:
Agency:	Agency:
Contact Info:	Contact Info:

Date of Referral:	Delivered via: ◯ Phone (emergency only) ◯ E-mail (encrypted) ◯ In Person (sealed envelope)
Any contact or referral restrictions?	◯ No ◯ Yes (specify any or all):

CCD mandatory data fields

Mandatory Data Fields for Collection and Sharing

1. ***first_name***
   a. String of characters
2. ***family-name***
   a. String of characters
3. ***gender***
   a. Male/Female/Other
   b. Captured as multiple choice
4. ***date_of_birth***
   a. Stored in the International date format YYYY-MM-DD; however can be captured as a number (i.e. 34)
   b. (if unknown, 1 Jan 1900 for adult, 1 Jan 2010 for child)
   c. Numerical string
5. ***place_of_birth***
   a. If in country of operations, captured by choosing from standardised drop-down list as determined by country consortium
   b. If born outside of country of operations, chosen from **drop-down** list of countries
6. ***community_id***
   a. Current location(community) person(s) resides. These can be permanent settlements like villages and hamlets or temporary housing establishments like refugee camps and temporary shelters for survivors of disasters
   b. Captured by choosing from standardised **drop-down** list as determined by country consortium
7. ***mobile_phone_id***
   a. Mobile phone number of person
   b. Captured as a numerical string with no country code
8. ***hh_size***
   a. Number of people living in same location
   b. Captured as a numerical string
9. ***Government issued identification document (if applicable)***
   a. Need to capture both type and ID number
      i. gov_id_type: Select from **drop-down** lists the types of government-issued ID that are applicable for the country (national identity card, passport, etc.)
      ii. gov_id: identification number of ID (i.e. passport number) captured as alphanumeric string

Meta Data Fields for Sharing

1. ***registering_org***
   - Name of organization creating original record of person
2. ***timestamp_orginal***
   - Time/Date record set created
3. ***modifying_org***
   - Name of organization modifying record
4. ***timestamp_mod***
   - Time/Date record set modified
5. ***staff_mod***
   - Name of staff member modifying record
6. ***reason_mod***
   - Reason for modification

Retained Mandatory Data Fields (not shared)

1. ***consent_to_capture***
   - Indicates if this person has given consent for their personal details to be captured and stored by organization
2. ***consent_to_share_info***
   - Indicates if this person has given consent for their personal details to be shared with other agencies.

Retained Meta Data Fields (not shared)

– staff_reg

- Name of staff member who is the original data collector

– staff_id

- ID number of staff member

UN MINIMUM CORE ASSISTANCE DELIVERY DATASET FOR AFFECTED POPULATIONS

MINIMUM DATA FOR HOUSEHOLD RECORD	
The below fields are for a household record set. Fields with (*) are minimum requirements for individual (not household) records. *Where a field is only required for individual-level assistance and not household assistance, that is specified.*	
FIELD LABEL	**FIELD DESCRIPTION AND GUIDANCE**
Informed consent (include purpose)*	With whom and for what purposes the beneficiary agreed to share personal data and should include the concept of onward transfer for the purpose of assistance delivery.
Source case/household ID*	ID from source organization (from which organization the data is being shared)
Original enrolment organization case/ household ID*	For deduplication and verification purposes, shows entity that conducted the original enrolment of the household; this may or may not be the same identifier as the 'source case/household ID'. This should always be filled out.
Original enrolment organization name	Name of the original enrolment organization
Size of household	Number of people in household
Case/household location/address*	Geographic fields should use IASC Common Operational Datasets where available; agreement at country level should be made about structure and format of address fields and what level of precision is required.
Full name of head of household*	Format of name fields should be agreed at country level.
Alternative assistance collector(s)*	Format of name fields and number of alternative assistance collectors should be agreed at country level.
Individual ID number of head of household*	ID from source organization
Individual ID number of alternative assistance collector	ID from source organization
Individual ID number of person to receive assistance* (only for individual assistance)	ID from source organization
Relationship to head of household* (only for individual assistance)	
Sex of head of household (sex of individual in case of individual assistance*)	
Sex of alternative assistance collector*	
Date of birth of head of household (date of birth of individual in case of individual assistance*)	
Date of birth of alternative assistance collector*	

Disability indicator for head of household (disability indicator for individual in case of individual assistance*)	Indication if head of household is disabled (in order to facilitate necessary support for them to access assistance); format and content of disability indicator to be agreed on at country level.
Disability indicator for alternative assistance collector*	Indication if alternative assistance collector is disabled (in order to facilitate necessary support for them to access assistance); format and content of disability indicator to be agreed on at country level.
Sex and date of birth of each individual within the assisted household*	Date of birth or year of birth when date unavailable
Sex/age cohort of the household members	Females age 0–4, Females age 5–11, females age 12–17, females age 18–59, females >60, total number of males, males age 0–4, males age 5–11, males age 12–17, males age 18–59, males >60, total number of males, total group size
Humanitarian profile category*	Indicates the IASC Humanitarian Profile Common Operational Dataset category for the beneficiary (i.e. Displaced [IDP, Refugee or Asylum Seeker, Others of Concern] or Non-displaced [Host, Non-Host]) https://www.alnap.org/help-libary/iasc-guidelines-on-the-humanitarian- profile-common-operational-dataset
Government ID or other civil society ID type and number*	Acceptable forms of ID to be agreed at country level; type should be a descriptor of what the ID number is.

INDIVIDUAL/HOUSEHOLD METADATA	
Date last updated	Date when record was last modified
Data collection methodology	Registration, community enrolment
Date originally registered	Date when original registration record was created

OPTIONAL FIELDS	
Biometric data of head of household and alternative assistance collector	
Photo images of head of household and alternative assistance collector	When another form of biometric data is not being provided
Other household IDs from other assistance agencies/organizations	Includes ID and ID description indicting source of ID
Phone number/account number of head of household and/or alternative assistance collector	For use where required for cash-based interventions (CBI) or other operational reason
Account number type	Description for type of account number above (e.g. bank card, cash card, bank account number, wallet, mobile money)
Account issue date	Start date/activation date of above account number
Account expiration date	End date/deactivation date of above account number

Jordan Red Crescent Society Ajloun president with IFRC Cash delegate introduces an agreement to a Syrian refugee. (Photo Credit: IFRC).

THE FUNDAMENTAL PRINCIPLES OF THE INTERNATIONAL RED CROSS AND RED CRESCENT MOVEMENT

Humanity

The International Red Cross and Red Crescent Movement, born of a desire to bring assistance without discrimination to the wounded on the battlefield, endeavours, in its international and national capacity, to prevent and alleviate human suffering wherever it may be found. Its purpose is to protect life and health and to ensure respect for the human being. It promotes mutual understanding, friendship, cooperation and lasting peace amongst all peoples.

Impartiality

It makes no discrimination as to nationality, race, religious beliefs, class or political opinions. It endeavours to relieve the suffering of individuals, being guided solely by their needs, and to give priority to the most urgent cases of distress.

Neutrality

In order to enjoy the confidence of all, the Movement may not take sides in hostilities or engage at any time in controversies of a political, racial, religious or ideological nature.

Independence

The Movement is independent. The National Societies, while auxiliaries in the humanitarian services of their governments and subject to the laws of their respective countries, must always maintain their autonomy so that they may be able at all times to act in accordance with the principles of the Movement.

Voluntary service

It is a voluntary relief movement not prompted in any manner by desire for gain.

Unity

There can be only one Red Cross or Red Crescent Society in any one country. It must be open to all. It must carry on its humanitarian work throughout its territory.

Universality

The International Red Cross and Red Crescent Movement, in which all societies have equal status and share equal responsibilities and duties in helping each other, is worldwide.

The International Federation of Red Cross and Red Crescent Societies (IFRC) is the world's largest humanitarian network, with 192 National Red Cross and Red Crescent Societies and around 14 million volunteers. Our volunteers are present in communities before, during and after a crisis or disaster. We work in the most hard to reach and complex settings in the world, saving lives and promoting human dignity. We support communities to become stronger and more resilient places where people can live safe and healthy lives, and have opportunities to thrive.

Follow us:

www.ifrc.org | twitter.com/ifrc | facebook.com/ifrc | instagram.com/ifrc | youtube.com/user/ifrc | tiktok.com/@ifrc