Guidance and best practices

HUMANITARIAN Data Protection

Innovative technologies and new environments

Guidance for IOM staff

Draft – July 2017

Acknowledgements

Authors and contributors: Amelia Stoenescu, Debora Gonzales Tejero, Garrett Flint, Helen Seeger, Mohamed Rizki, Rutendo Urenje, Sebastian Ancavil, and Stephanie Daviot, Vlatko Avramovski; Coordination: Nuno Nunes

Special thanks to IOM staff who responded to the call for information by completing the humanitarian data survey. More than 180 IOM staff in 31 IOM offices who gave their time to fill in the surveys, to provide information on their work with data and in identifying specific experiences when questions or issues were raised. Their input and experiences were essential in shaping the scope and the content of these guidelines (in alphabetical order): Aaron Manyumbu, Abdoulaye Djibo, Abdul Ahadi, Abdul Nasir Kakar, AbedElQader Garaibeh, Ahmad Maiwand Ahadi, Ahmad Shoaib Almas, Akinyera Aanuoluwapo, Alberto Preato, Alejandro Galarza, Alhadj Mahamat Ali, Amal Amroussi, Ameni Yaruta, Ana Ruth Reyes, Andy Armansyah, Antonio Meono, Asma Saeed, Aswad Muhammad, Badel Khan Awan, Badel Khan Awan, Balazs Lehel, Bekim Ajdini, Ben Christopher Mbaura, Betty Asienzo, Bintou Sacko, Bogdan Danila, Camaldyne Alhoussene Touré, Carmela Godeau, Carol Lida Sasa, Carolina Brill, Ceren Dönmez, Claudia Carolina López Taks, Claudia Costa Pereira, Claudia de los Ángeles López, Daniel Bagameri, Daniel Salmon, David Barreno, David Coomber, David Morales, Dembele Bambo, Dessalegn Gurmessa, Dina Afzali, Diwakar K.C, Djibrine Mahamat, Dr Rama Jayavanth, Dula Wakassa Duguma, Duncan Sullivan, Elda Goci, Elena Gudalupe Grimaldi, Ester T Gigir, Fawaz Ahmed Ali Asemi, Federica Sanguinetti, Gamuchirai Perekwa, Gideon Madera, Giovanni Cassani, Giulia Musati, Guillermo García, Gul Muhammad Ahmadi, Hasballah, Hisham Elawad, Idriss Achour, Ihsan Hussein, Ivona Zakoska, Jacobo Muelas, Jeff Labovitz, John Chaduka, Jorge Peraza-Breedy, José Dilber Gómez Hernández, Joselito Cabaña, Josephine Mann, Josiah Ogina, Kashif Khan, Kathy Pinargote, Kendra Rinas, Kevin Merkelz, Khalil Omarshah, Khaoula Ouederni, Kiki Damayanti, Kojo Annet Kenyi, Kristina Uzelac, Kyaw Kyaw Tun, Kyaw Soe Khine, Lalu Atikdar Firman Hakim, Laura Armstrong, Laura Bartolini, Laurence Hart, Ligia Yanis Pérez Solís, Lorelle Yuen, Lumanyi Michael John Batista, Ma Aye May, Magham Ibrahim Bakhiet, Mahamadou Seyba Traore, Mahamat Akouya, Mahwish Irfan, Makanjuola Adeola Goriola, Manahil Qureshi, Manisha Rai, Marcello Goletti, María Cristina Castillo, Marie Karleskind, Marina Schramm, Mario Lito Malanca, Masood, Mazen Aboulhosn, Mehari Hagos, Melissa Tui, Michael Jerry Nwune, Michelle Kajoka, Mingizem Maru, Modher Alhamadani, Mogga Julious Alex, Mohammad Omar Majeedi, Mohammed Alkali Tayib, Mohammed Dizaeey, Mohamud Omer Abdi, Monther Numan Kaid Saif, Mouhammad Hazzouri, Muath fadl mohamad al mahdi, Muhammad Awais, Muhammad Madya Akbar, Muhammad Omar Hashimi, Munir Ali, Nabin Sharma, Nadine Nkubonage R., Navin Karki, Nenette Motus, Olivier Balagizi, Ousmane Kore, Ousmane Yoroté, Patricia Schiavinato, Peter Murorera, Rachel Dolo, Radheshyam Krishna KC, Raja Ram Dhungana , Rangarirayi Tigere, Ranjan Rijal, Richard Danziger, Riju Stephen, Robel Kasaa, Rodas Yeheyes, Saleh Bin Selm, Salvatore Sortino, Sanjaya Deshar, Sarah Craggs, Sarfaraz, Sarfraz Sargand, Sasleem Taha Aljarmouzi, Sebastian Berkovich, Sergio Ovalle, Shahzad Ahmad Mamoon, Sharon Nerius, Silvio Majorino, Simon Kafu, Solomon Buka, Stefanie Briones, Stuart Simpson, Sujan Paudel, Suman Karmacharya, Sunita Mainali, Susan Atala, Talingapua Quentin, Tatjana Cerovina, Temitope Omoyemi, Thomas Mek, Tjossy Mitaart Sipasulta, Tun Moe Khaing, Umber Gaho, Víctor García, Wakene Tadesse, Waqas Ahmed, Yaseen Al Shereda, Yitna Getachew , Yohannes Mulu, Zainullah Sultani, Zoe Nasim Akhter

With special thanks to Jos Berens, Stuart Campo, Nathaniel Raymond and Linnet Taylor who contributed their knowledge and experience to these guidelines.

Special contributions:

Review: Raul Soto, Nikki Herwanger, Christina Vasala Kokkinaki

This work would not be possible without the contribution of the European Commission, Netherlands and …

Guidance and best practices in humanitarian data protection Innovative technologies and new environments

(Guidance for IOM Staff)

Contents

Acknowledgements 2

Contents 4

Organizational process 6

Background 7

Purpose of this document 8

Ethical data management 10

Data Processing cycle 10

Design 10

Collection 12

Analysis 12

Storage 14

Sharing and Third Party Usage 14

Threats to IT infrastructure 15

Recommendations 15

Innovative Technologies 18

Remote sensing 19

Satellite imagery 19

Design 19

Collection 20

Analysis 21

Storage 22

Sharing 22

Unmanned aerial vehicles (UAVs) 22

Design 23

Collection 24

Analysis 25

Storage 25

Sharing 26

Crowdsourcing 26

General recommendations 28
Social media 29
Design 31
Collection 32
Analysis 32
Storage 33
Sharing and use 34
Mobile Networks 34
Design 36
Collection 37
Analysis 38
Storage 39
Sharing 39
Instant messaging applications 39
Design 40
Collection 42
Analysis 42
Storage 42
Sharing 43
Advanced computing 44
Volunteer networks 44
Advanced data analytics 46
General recommendations 46
Phenomena that shape risk in advanced analytics: 47
Annexes 48
Annex 1 Existing data protection guidance from IOM and other actors 49
General data protection guidance 49
Satellite Imagery 51
SMS and mobile telephones 51
Messaging Apps 52
Social media and networks 52
Unmanned Aerial Vehicles (UAV) 52
Additional guidance 53
Annex 2 Staff Survey 54
Copy of survey questionnaire 58
Annex 3 DTM Data Protection Checklists 67
Annex 4 Quality Assurance Processes for Displacement Tracking Matrix products 71

Organizational process

This guidance document was prompted by requests from IOM field staff for support to their work with data in humanitarian contexts. While data protection publications and resources have been produced by many organizations, including IOM's Data Protection Manual, staff in IOM offices worldwide requested further guidelines covering innovative technologies and new environments that are sometimes not explicitly mentioned in existing guidance. To respond to this request for further support, IOM's Department of Operations and Emergencies (DOE) led the production of these guidelines in coordination with the Legal Department (LEG), and Information and Communication Technology (ICT) Division. Representatives from each department worked together to agree on the contents and scope of the final document. This cross-departmental collaboration ensured that different aspects of data protection were covered, including operational aspects, legal aspects and technology.

The production of these guidelines involved a review of current data protection literature and policies, a survey of IOM staff, and input from experts and practitioners in the field of data and technology.

The literature review focused on existing guidance for data protection as produced by organizations including IOM, guidance on the use of innovative technologies in humanitarian settings and other relevant publications. The complete list of consulted guidance can be found at the end of the document. The literature review identified gaps and contributed to the structure of the document, it also provided the framework for the staff survey that gathered information about real-life experiences faced by IOM staff working with data.

A survey was developed in order to identify specific ethical issues or questions faced by IOM staff. The survey began by establishing the profiles of the respondents. Survey participants were asked about their mission, their job, and how their work relates to data. Survey responses were then analysed according to the profiles, in order to understand the variety of experiences and needs. The survey also asked additional questions to those who reported facing challenges or ethical issues in their work with data, in order to gather specific examples.

IOM receives many requests for additional guidance on data protection from staff all over the world. For the purposes of the survey, IOM staff currently working with the Displacement Tracking Matrix (DTM) were chosen as the target respondents for the survey. DTM coordinators were chosen from 23 IOM offices from around the world. Each coordinator was asked to conduct the survey with as many staff as possible. Working through the DTM coordinators ensured the office's ownership of the project and contributed to a high response rate. After receiving responses from 155 DTM practitioners working in 27 countries and comprehensive feedback, it became clear that this exercise was warmly welcomed by IOM offices. It was decided to roll out a second round of the survey, this time focusing on senior management, Chiefs of Mission, and Regional Directors. In total, 180 staff responded, providing strong evidence for the need for further data protection guidelines as well as rich individual examples of challenges staff face when working with data. The survey results and personal experiences furnished these guidelines with both best practices and lessons learned. A detailed description of the survey as well as analysis of the results is available toward the end of this document. The survey questionnaire is available in Annex 2.

After incorporating the survey results, experts from different organizations with experience in various aspects of data provided their input. Experts were identified among existing contacts from ongoing discussions around data management. Each expert was individually asked to review the draft and provide input.

On completion of the draft, the document was shared with colleagues from three IOM departments: DOE, ICT, and LEG. This was not a formal endorsement process, as the document is not intended as a mandatory instruction, but rather a list of recommendations and best practices. The review ensured that the document followed existing standards and procedures, aligning to other data-related IOM policies.

The resulting guidance document *Guidance and best practices in humanitarian data protection* is intended for internal use by IOM staff working with data in emergency situations, but is also made available to the public.

Background

Datafication is a modern technological trend turning many aspects of our life into computerised data and transforming this information into new forms of value. Examples of datafication as applied to social and communication media are how Twitter datafies stray thoughts or datafication of HR by LinkedIn and others. (Data from: Wikipedia)

Humanitarian response is not yet data-driven, but it produces large volumes of data. Information is essential for designing and planning response, prepare distributions and assistance, and to monitor or evaluate impact. While questionable whether humanitarian actors are prepared to make full use of data produced in emergencies, the use of innovative technologies has revolutionized the way humanitarian actors interact with data and population needs in a way never before possible

With changes in technology and greater access to information, there are more risks, requiring humanitarian actors to responsibly handle humanitarian data. Information management has become an integral part of the tasks of many IOM staff who deal directly with data every day. With the increased exposure to information, specifically the collecting and sharing data about the world's most vulnerable communities, there are increased ethical implications and responsibilities for IOM staff. IOM has been a leader in its proactive approach to data protection, being among the first international organizations to provide its staff with a data protection manual and principles. While these principles do exist, there have been numerous requests for additional guidance from IOM staff to respect IOM's data protection principles while working with innovative technologies that aren't explicitly mentioned in existing guidance.

It is essential that IOM staff handle humanitarian data and use technology responsibly, always focusing on the protection the rights of the affected populations. Improper collection, handling and sharing of data on affected populations can create very real risks that may threaten the right to security and privacy of individuals, communities, or specific demographics. For several reasons, risks may become even more pronounced with the use of innovative technologies. For example, without focusing on the rights of the affected population, it can be tempting to use these technologies to gather more information than is needed, merely because it is possible to do so. This data will serve no purpose and can cause harm to the affected population and humanitarian response.

One risk comes from big data. Big data is widely described as heavily comprising of the "4 V's": Volume (huge amounts of data), Variety (different types of media from many sources), Velocity (produced extremely fast), and Value (high usage potential).[1] Though IOM does not produce anything in the realm of big data, mechanisms developed for big data analysis increase the risk of unintended analysis of IOM data sets when combined with other data. More information on this can be found in the *advanced data analytics section* in this document

Even if all the data collected using innovative technologies is verified to be accurate, it may still paint an incomplete picture. Differences in access to and participation with technologies varies widely based on several factors including age, gender and socio-economic status. Humanitarian action that is based on incomplete information may be just as harmful as action that is based on false or unverified information and can potentially have negative consequences for the affected population.

Additionally, it has long been known that humanitarian data that falls into the wrong hands, could be used to further exploit or harm individuals and communities: companies could use the information to create ethically questionable marketing strategies, governments could use the information to advance a damaging political agenda, or armed groups could use data to target specific persons or groups. Now, the threat of data ending up in the wrong hands is further exacerbated by the increased usage of innovative technologies that enable cyber-warfare, digital crimes, and government surveillance.[2]

Purpose of this document

While the title of this document refers to the protection of humanitarian data, it should also be interpreted as the ensuring that all humanitarian information activities[3] (HIAs) protect the rights and dignity of the affected population. This is especially important as IOM staff increasingly include the use of innovative technologies to conduct HIAs.

It is important that all work with data about affected populations places the rights of those individuals at the heart of any HIA. IOM staff will work to protect the rights of the affected people in regards to information, recognizing that these rights include: "*the rights to access, transmit and benefit from information, protection from harms that may result from the provision of information during crisis; expectation of privacy and data security; to have agency over how their data is collected and used; and*

[1] Whipkey, Katie & Verity, Andrej. *Guidance for Incorporating Big Data into Humanitarian Operations*. Digital Humanitarian Network. September 2015

[2] UNOCHA. Humanitarianism in the Age of Cyber-Warfare.

[3] "Activities and programs which may include the collection, storage, processing, analysis, further use, transmission, and public release of data and other forms of information by humanitarian actors and/or affected communities"(The Signal Code)

the ability seek redress and rectification when data pertaining to them causes harm or is inaccurate."[4] When working with humanitarian data, whether during the collection, analysis, storage, sharing or usage stages of the data lifecycle[5], IOM staff should prioritize upholding these rights. When working with personal data, it is mandatory for each staff member to know and apply the IOM Data Protection Manual and Principles, which are explained further with practical examples in IOM's Data Protection Manual.

This document provides guidance on working with innovative technologies that are not explicitly mentioned in the IOM Data Protection Manual. The guidance is presented under (1) **ethical data management**—covering guidance to be applied in general throughout the data management cycle when using innovative technology, regardless of the technology used, and (2) **innovative technologies** —containing guidance and best practices to be applied by IOM staff when using a specific technology to conduct humanitarian information activities, and (3) **advanced computing**—including information on innovative human and machine methods to process big data that is used for humanitarian purposes.

This document does not claim to be exhaustive or final; there are many technologies that are not covered, and surely many more that have yet to even be identified. Rather, it is a tool, meant to empower IOM staff to work with new data sources—using innovative technologies confidently and responsibly, to explore new partnerships, and to ensure that the right to timely, accurate, and consistent information is considered a basic need of an affected population. This guidance document will be revisited and updated to include new technologies as they evolve. The use of new methods and technologies by IOM staff should be communicated to HQ to be included in revisions of this document. Any IOM staff who decide to work or experiment with technologies not covered in this guidance document should be aware that this presents a potential threat and may cause serious harm to vulnerable populations.

[4] *The Signal Code*: *A Human Rights Approach to Information During Crisis.* Harvard Humanitarian Initiative (HHI). January 2017

[5] Design, collection, analysis, storing, sharing (including 3rd party usage.

Ethical data management

Many recognize the need for minimum technical standards when using ICTs. However, the development and advancement of technologies greatly outpace our ability to provide standards. This requires an ethical approach in providing guidance. [6] As technologies advance at a rapid pace, an ethical approach aims to equip staff to better manage risks presented by innovative technologies. Ethical data management requires staff to respect and care for all data throughout the data processing cycle and adopt a *do no harm* outlook.[7] This document has identified the data processing cycle as: collection, cleaning and verification, storage, analysis and sharing. It is mandatory that IOM staff manage data in a way that protects the privacy and security of the data subjects. Staff should realize that these risks go beyond just personal data to any data that can be used to identify a person, or members of a specific community or demographic.

When dealing with data that can be used to identify a person, community, or demographic, staff must apply the IOM data protection principles to all phases of the data processing cycle. These principles are: lawful and fair collection, specified and legitimate purpose, data quality, consent, transfer to third parties, confidentiality, access and transparency, data security, retention of personal data, and ownership of personal data. This section will integrate the data protection principles into all stages of the data processing cycle: including design, collection, storage, analysis, and sharing. This section is meant to be over-arching, giving guidance on the protection of privacy throughout the data processing cycle that apply across all technologies. Guidance for the processing of data collected through specific technologies is provided in the following section.

Data Processing cycle

Design

Before beginning any HIA, key factors regarding data protection should be identified for each stage of the project. Before beginning, staff should identify the data and level of detail required, the use of the data, with whom the data will be shared, and the office's capacity to process the data. The Data Protection Manual provides a set of checklists to be filled out before completing any data collecting exercises. The checklists include key considerations to be made in regards to: risk/benefit assessments, sensitivity assessments, consent, confidentiality, data security, data quality, and sharing personal data with third parties. These checklists have been adapted for DTM projects and are available in Annex 3 of this document. However, beyond the checklists, it is important to identify the risks and potential ethical concerns associated with collecting the data for a project as well as any possible implications that may be presented when using a specific technology.

Establish Capacity

More data increases the time, resources, and energy needed to process the information and adds to the risks to individual privacy and security throughout the data processing cycle. The capacity to process the information should also be a determining factor in the amount of data to be collected. One way of

[6] Martin-Shields, Charles. *The Technologist's Dilemma: Ethical Challenges of Using Crowdsourcing Technology in Conflict and Disaster-Affected Regions.* Georgetown Journal if International Affairs. Summer/Fall 2013. (pp. 157-163)

[7] UNOCHA. *Humanitarianism in the Network Age.* 2012

ensuring the capacity requirements are met is engaging with volunteer networks to provide specific tasks along the data-processing cycle. Consult the *volunteer networks* section of this document for further information.

Choose a technology

The first step in identifying the proper technology to be used is understanding the local communication and data ecosystem. This can be done by including in the needs assessment indicators to understand the local information channels, including the identification of trusted communication channels and technologies. HIAs should work with existing information channels in a given context, not duplicate or compete with them. It is also important to identify how information channels may be different based upon demographics. Activities that fail to take into account the varied access to different information channels will produced biased information. Activities should use a holistic approach, covering as many channels as necessary to obtain accurate information.

The local threats to information specific to a context, such as the level of government surveillance and digital crime groups should also factor into decisions in the design phase.

The use of a specific technology may be seen by members of the community and the government alike as a potential threat to their safety and security. This can be due to past experiences, or perceived risks and dangers associated with any of their information being made available to the wrong hands. Using an inappropriate technology that causes fear, discomfort or suspicion in community members or local authorities purely for the convenience of the data collection process is counter-productive, as responses are likely to be incomplete and potentially biased, while compromising perceptions of IOM, acceptance of programming and security of staff. Staff are encouraged to collaborate closely with local authorities before launching any humanitarian information activity.

Conduct assessments to establish local views of each technology. Local staff, communities and local authorities should be consulted regarding the history and perception of any particular technology. Technologies used for data collection such as tablets, GPS devices, other mobile devices, and UAVs may be viewed particularly negatively in conflict situations. These technologies may raise fears in the public that information on their location, or other potentially identifiable information can be used to further target them as individuals or as a community. Likewise, government or non-state actors often have fears of strategic or other dangerous information ending up in the wrong hands.

In an example given in the survey, one staff shared an experience where the decision for enumerators to use tablets during the data collection exercises was met by fierce opposition and distrust by the local authorities. Since the IOM office was operating in a conflict-affected location, the authorities were afraid that the tablets were being used to send data and coordinates to enemy groups. In some cases, this even led to short detainments of enumerators who were found using the tablets. Once the IOM office was aware of the government's view of the devices, plans were altered to exclude the use of any mobile devices in the data collection process.

Conduct risk-benefit assessments

It is important to identify the risks and benefits of using any particular technology. While technology facilitates many aspects of IOM's work, it is important to identify any possible risks associated with using that technology. Improper or unethical data management practices lead to risks for IOM as an institution and harm to the populations the organization aims to serve. For each of the technologies in this document, it is important to carefully weigh the risks and identify whether or not the risks can be appropriately mitigated, and whether the benefits outweigh the risks.

Collection

Lawful and fair collection

The Data Protection Principles assert that collection of personal data must be lawful and fair. Data subjects must understand all the uses of their data and provide their consent accordingly. In addition, the Principles state that data collected must serve a specified and legitimate purpose. Collection activities should be limited to those data that are necessary and will be used for humanitarian actions. Data collection efforts that go beyond a specified purpose should be discontinued. Innovative technologies are subject to the same principles, however they also present new challenges.

Managing expectations

New technologies can facilitate the interaction between humanitarian actors and the affected population, but can also increase expectations of the affected population. Once data is collected, it brings with it the ethical obligation to respond to the identified needs; thus collection activities should be limited to the capacity to respond. It is essential that expectations are well managed at the beginning of data collection activities.

Conversely, some technologies limit the interaction between humanitarian actors and affected populations, increasing the difficulty to obtain consent. For example, data may be collected using remote methods, or obtained through datasets from governments or the private sector. Such data will likely not have been provided with the consent, or even the awareness of the data subject. It is essential that steps be taken to limit the amount of personal data collected regardless of the technology used. Guidance will be given in regards to particular technologies in the subsequent chapters of this document.

Compatibility of data

Data should be collected in formats compatible with those used by other humanitarian and development organizations. Efforts to standardize data will decrease the effort required for collaborators in the data ecosystem to use data collected by IOM.

Analysis

Cleaning and validation

The first and essential steps in analysis are cleaning and validating the data. These include detecting and correcting corrupt, inaccurate, or incomplete records from data sets.[8] Information that is incorrect and

[8] Wikipedia.org

incomplete can lead to inaccurate or miscalculated responses, which have the potential to cause further harm. All data, whether gathered by IOM, or provided by a third party, should undergo cleaning and validation. Many of the surveyed IOM staff shared experiences of data about affected populations being skewed whether intentionally or inadvertently. It is vital that these discrepancies are adjusted before any actions are taken by IOM or other actors. No matter the technology used, measures should be taken to assure data quality. For DTM staff, the Quality Assurance Process flowchart (available in Annex 4) should be followed throughout the phases of the data life-cycle.

No matter the technology used, data will often come from multiple sources and in multiple formats. It is important to ensure there is adequate capacity to process the data coming in from any HIA. Results from any data collection activity should be cross-checked whenever possible against other sources.

Concerning personal data one of IOM's Data Protection Principles is data quality, and the manual provides specific guidance on how to ensure high quality of personal data.

Triangulation

Whenever possible, data should be checked against other sources. Cross verifying information can identify anomalies and discrepancies and ensure the data more accurately reflects reality. Data triangulation usually refers to three approaches: data source triangulation, methodology triangulation, and theory triangulation.

Data Source Triangulation: Using a variety of sources including both primary and secondary data.

Methodology Triangulation: Combining multiple methods, tools and technologies to gather data.

Theory Triangulation**:** Analysing data using more than one theoretical approach to interpret the data

Conducting analyses

Once the data has been cleaned and validated, analyses may be performed to turn the data into actionable information. Analyses should be limited to datasets that do not include any sensitive or confidential information, such as data that is personally or demographically identifiable, or locations of critical infrastructures in conflict zones. When dealing specifically with personal data, the confidentiality of individuals or communities must be maintained during analysis, as outlined in the IOM Data Protection Manual.

Protection specialist advice should be sought when analysing and publishing data on protection concerns including unaccompanied children (UAC), gender based violence (GBV) or the prevention of sexual exploitation and abuse (PSEA). This data while non-personal in character has an element of sensitivity which, if submitted to the risk-benefit assessment stated in the Manual, would most likely incur more risks than benefits for the data subjects.

Anonymization of data

IOM staff should ensure an adequate level of anonymization through an effective process to ensure sufficient protection and privacy. To anonymize data correctly, staff must first understand which data requires anonymization. This data includes, among others, the names of individuals, their ethnic and racial origin, political and religious beliefs, sexual orientation, physical and/or mental health. Some anonymization techniques include: data coding, pseudonymization, or simply removing sensitive data fields.[9] With advanced data analytics tools and immense datasets, complete anonymization becomes extremely difficult. Datasets that do not appear to have any personal information can be combined with others to draw accurate conclusions about the identity of the data subject. It is important to think, not only about the data contained in one dataset in isolation, but the total data available.

Partnerships for data analysis

To streamline the analysis process, it may be recommended to establish a partnership with volunteer networks or even companies in the private sector that provide such services. However, this introduces another layer of risk to data protection. Thus, any such collaborations should be clearly outlined in a written agreement with the private sector partner, which will also include an obligation for the partner to follow IOM's Data Protection Principles. Guidance for establishing partnerships with companies and volunteer networks is given in the relevant chapters of this document.

Storage

All information should be handled with care, taking proper measures to avoid unauthorized modification, unlawful destruction, accidental loss and improper disclosure. Any individual who provides information reserves the right, for example, to rectify or delete their information at any time and storage practices should therefore allow for such instances. As with data quality, proper storage techniques and guidelines for personal information can be found in the IOM Data Protection Principles and Manual.

The storage and handling of data depends on its classification. Classification should follow the guidelines provided in the Migration Data Governance Policy (IN/253)[10]. Electronic records should be classified per their level of sensitivity and they should be clearly marked prior to transferring it to authorized persons or saving it to electronic storage areas. Access to electronic records should be limited only to authorized users. The guidelines in the Data Protection Manual should be followed when classifying electronic records containing personal data.

Sharing and Third Party Usage

Data collected by IOM may be valuable to other organizations, including but not limited to: international organizations (IOs), non-governmental organizations (NGOs), media[11], and governments. Making humanitarian data accessible is one of the core tasks of humanitarian actors. There are many initiatives to establish open platforms of humanitarian data, allowing others to review, analyse and use the information. Data that is shared and combined becomes more complete, accurate, and useful. However, making data more accessible carries many risks. Those who have access to the data may use it for activities

[9] IOM Data Protection Principle 10

[10] https://dmsportal/PublishedDocuments/Instructions/Migration%20Data%20Governance%20Policy%20Instruction%201st%20May%202017.pdf#search=migration%20data%20governance%20policy (only available through IOM intranet)

[11] Note that any request from the media to IOM staff for information should be coordinated with MCD.

that go against IOM’s mandate, or against the specified purpose for which the information was gathered. This means that throughout the data processing cycle, IOM staff should keep potential secondary uses of the data in mind, particularly during the collection and dissemination phases. IOM staff should consult the IOM Migration Data Governance Policy[12] for guidance on the classification and sharing of data. In conflict situations, data should not be made public, even if the data are non-personal.

Before sharing any data, a written agreement should be in place between IOM and the entity requesting the information. Data sharing agreements and partnerships should be reviewed by LEG and refer to IOM Data Protection Principles in order, for example, to protect the analysis process and prevent unintended uses. The analysis and sharing of data will consider the humanitarian principles of humanity, neutrality, impartiality and independence, as well as risks of publishing non-personal sensitive data.

The sharing of personal data should comply with the IOM Data Protection Principles and LEG should always be consulted for the written agreement on data sharing that must be in place.

Threats to IT infrastructure

Throughout the data processing cycle, IOM staff should ensure the security of their IT infrastructure. There are many groups that may try to target the IT infrastructure to obtain humanitarian data. Hacking systems to retrieve personal information, deliberate attacks on an organizations IT security infrastructures, and countless other activities place humanitarian data at risk. In the staff survey, one IOM office reported cited examples of emails and websites from UN offices in the country that were hacked in the past. Cyber-attacks may be carried out for many motivations including targeting humanitarian organizations, identifying the recipients of aid, or targeting at risk or marginalized groups. Beyond hackers and online-groups, there is a growing risk of nation-state attacks to IT infrastructures. In his keynote address at the RSA Conference 2017, Brad Smith, President of Microsoft Corporation, stated, “The past year [2016] has witnessed not just the growth of cyber-crime, but a proliferation in cyber-attacks that is both new and disconcerting. This has included not only cyber-attacks mounted for financial gain, but new nation-state attacks as well.”[13]

Recommendations

Given the constant development of cyber-threats, it is important to understand how to proactively maintain the integrity of the IT infrastructure. To provide direction and guidance on the principles to be applied for maintaining the confidentiality, integrity and availability of IT resources throughout the organization, the ICT department compiled a list of policies and guidelines under IN/123[14]. This document pictures a risk based approach on what regards security. It means that IT solutions must be assessed from a business impact perspective and then, based on this impact, the right mitigation controls need to be put in place. It requires close collaboration with the ICT team.

[12] IN/253 only available through IOM intranet portal

[13] Smith, Brad. *The Need for a Digital Geneva Convention*. RSA Keynote Address. San Francisco CA. 14 February 2017

[14] Only available through IOM intranet portal.

Below is a brief list of personal precautions all staff should adopt.

Vulnerability patching

To reduce the thread of vulnerabilities being explored it is important to patch them as soon as they are discovered. Incorporate server and workstation patching as part of the weekly/monthly regular IT activities. Turn on the Automatic Updates feature in your Windows operating system. You will be notified of any updates before downloading them or automatically installing them. Mobile devices (phone, tablets, etc.) must also be patched on a timely basis.

Operating systems

Always use the latest available version of software. In the last few years vendors had put a lot of effort on including new security features in their products. It is also applicable for legacy system that must be migrated to new versions as soon as possible. It will reduce the IOM's exposure to security vulnerabilities but also reduce the effort of maintain different versions of the same software.

Encryption

Always encrypt data at rest. During a field mission the risk of IT devices, either personal or belonging to IOM, getting lost or stolen is always present. To make sure that the information contained on these devices cannot be accessed by external parties, always use encryption. For laptop and workstations, use Bitlocker on Windows devices and FileVault on Mac. For USB storage devices, please use Bitlocker to Go.

Mobile Devices security

As stated above, IT devices can be lost or stolen. On what regards mobile devices it is important that the impact of these events is mitigated.

Do not store any data that you cannot afford losing on mobile devices. Or any data that if they fall into the wrong hands, there can be severe consequences for IOM beneficiaries and IOM. Consider using cloud services to store information. IOM is currently implementing OneDrive for business and this can be used to store all your files while on the go. In top of that, it is strongly recommended to also backup your device at least every month.

While accessing IOM's information please use AirWatch. It allows you to access your e-mail and IOM's internal resources in a secure way. Follow the guidelines given in IN/76 Mobile Device Guidelines for Field and Headquarters, available through the IOM intranet portal.

Set a passcode for your mobile device. A minimum of 8 characters' passcode is recommended. Where possible, the passcode should be a combination of numbers, letters, punctuation, and special characters.

Mobile devices that are used for data collection, such as tablets, also involve risks, including improper or unauthorized access to the data in storage or during transmission to IOM servers. Mobile devices can store geographic locations (via GPS) thus potentially identifying individuals or target populations. Staff should carefully weigh the risks and benefits of using such technology, as well as understand the local perceptions of such technology, whether among the community or the government. Staff should always consult with the local government before using mobile devices for data collection. Ensure that data storage and/or transmission procedures have an adequate level of protection, such as encryption, passwords, and using secure Wi-Fi links.

For further IT specifications, please consult “ICT Standards and Guidelines IN/88”.[15]

Wireless security

Connecting to open, non-password protected Wi-Fi and internet connections, particularly in locations that are publicly accessible and known to be used by expatriates is risky. Before connecting to a public wireless network make sure it is a legitimate network. Don’t assume that a hot spot is real based on its name. If you’re at a hotel or café, ask a manager to confirm the name of its network—and that the network is encrypted (i.e., locked and password protected). While online, stay on encrypted channels by using the website prefix https (rather than http). It is always recommended to use IOM’s VPN client (Cisco AnyConnect) to establish a secure connection between your computer and IOM infrastructure.

[15] https://dmsportal/PublishedDocuments/Instructions/IN88%20ICT%20Standards%202016%20Rev1.pdf#search=ICT%20STANDARDS (only available through IOM intranet)

Innovative Technologies

Remote sensing

Remote sensing allows the gathering of information from a far distance or high altitude. Remote sensing technologies have been used in the past for military and agricultural purposes, and more recently, for humanitarian action.[16] Satellites and unmanned aerial vehicles (UAVs) help humanitarians collect information in areas that are inaccessible, or so large, that remote sensing is the only feasible means to gather such vast amounts of information. This document provides guidance on using satellite imagery and UAVs as remote sensing tools.

Satellite imagery

Satellite maps and Geographic Information Systems (GIS) are used increasingly in the response to emergencies. Satellite imagery can be used to provide up-to-date information for assisting displaced populations. The uses of satellite imagery can include the mapping of shelters and structures in camps, detection of displaced populations, supporting site selection, conducting rapid assessments, and documenting the impact of disasters.[17] One famous example is the use of satellite imagery to capture information on the intentional targeting of civilians in conflicts in Central African Republic, Sudan, and South Sudan. Humanitarian actors could study high resolution satellite images to analyse damage to traditional civilian dwellings, known as *tukuls*, as an accurate indicator for attacks led by armed groups against civilians.[18]

Using satellite imagery has many benefits for humanitarian response, but does not come without certain risks to the privacy and security of affected communities. Images could potentially be used to identify and target individuals, communities, or critical infrastructures.

This section provides best practices in managing information received or requested from satellite imagery to help mitigate these risks and safeguard the rights of the affected population. The section includes guidelines for managing satellite imagery data as well as key points on critical infrastructure, real time satellite imagery, posting requests in volunteer networks and use of imagery in conflict situations.

Design

A scale of necessity should guide the decision whether to request and use satellite imagery. Staff should balance the urgency of the need, the risk of delays and additional data confidentiality risks related to engaging external partners and the quality of analysis that can be performed in house. Before requesting satellite imagery, staff should first establish: the purpose of the project, which objects will be identified, the limitations of the available satellite imagery, and how the data will be processed, stored and shared. In natural disasters, staff should coordinate early on with the national geospatial or geographical center, or statistics office in its absence.

[16] Omara, Samuel. *How Remote Sensing is a Game-Changer in Humanitarian response*. 31 May 2016. https://www.linkedin.com/pulse/satellite-imagery-game-changer-humanitarian-response-samuel-omara

[17] *Satellite Imagery Interpretation Guide: displaced population camps. Harvard Humanitarian Initiative* http://hhi.harvard.edu/sites/default/files/publications/satellite_imagery_interpretation_guide.pdf

[18] *Satellite Imagery Interpretation Guide: intentional burning of tukuls.* Harvard Humanitarian Initiative. http://hhi.harvard.edu/sites/default/files/publications/siig_ii_burned_tukuls_3.pdf

Collection

Publicly available imagery

Many websites provide satellite imagery for download at no cost to the user. These images are publicly available and can be used to create maps and perform other GIS activities.

Digital Globe

High resolution satellite imagery to support disaster recovery. Releases pre-event imagery, post-event imagery and crowdsourced damage assessments. https://www.digitalglobe.com/opendata

International Charter Space and Major Disasters

The Charter is the result of UNISPACE Conference III held in Vienna in 1999. Agencies who joined the charter aim to unify the collection and sharing of satellite imagery to support response to natural and man-made disasters. The Charter has an agreement with UN OOSA and UNITAR/UNOSAT to provide support to UN agencies. UN OOSA and UNITAR/UNOSAT may submit requests on behalf of users for the United Nations. https://www.disasterscharter.org/web/guest/about-the-charter

EarthExplorer (USGS)

Used to obtain historical data for many purposes. This site requires registration. After filling a questionnaire for approval, more detailed imagery can be updated and used for emergencies.
https://earthexplorer.usgs.gov/

Open Aerial Map

A branch of Open Street Map, uses primarily drone images, but has satellite imagery available. https://openaerialmap.org/

Copernicus Open Access Satellite Hub

Copernicus is fed by a set of satellite “sentinels” focusing on six thematic areas: atmosphere, marine, land, climate, emergency and security. The system requires a login.
http://www.copernicus.eu/main/copernicus-brief (Sentinal)
https://sentinel.esa.int/web/sentinel/about-sentinel-online

Imagery during open conflict

Imagery in areas of conflict should be requested through specialized agencies such as UNOSAT, in coordination with IOM HQ. The global level the focal point for imagery is responsible for contact with specialized agencies such as UNOSAT, and can be reached at DTMGlobalImagery@iom.int. Coordinate with partners and include them before making the request, as needed, to reduce duplication and streamline the collection activities in an emergency.

When preparing to request satellite imagery, please identify:

- Geographic bounds
- Temporal bounds
- Operation the imagery will support
- How the data will be used
- A specific question

- All partners who will use the data

Do not use a large area of imagery; if it is of no use or specific interest for the project, it is best to keep the area as small as possible. The request should specify to omit critical infrastructure, such as military facilities, power plants, and schools, especially in the context of conflicts.

Real and near-real time satellite imagery

When real time satellite imagery is available for an area affected by an emergency, coordinated discussions should be held with relevant actors, including governments, regarding whether or not to use it. Real and near real time satellite imagery should still be subject to the same analysis standards. Real time satellite imagery is not currently publicly available, but the necessary technology is under development and such imagery may become publicly available in the near future.

Analysis

Cleaning and validation

In conflict locations, any imagery that allows for the identification of individuals or groups, or contains data classified as confidential (e.g. location of a hospital in a conflict zone) must not be used. Staff should only use satellite imagery that has been cleaned and modified by specialized agencies such as UNOSAT, instead of raw images. Raw imagery has not been modified in any way and could create a liability for IOM by showing more information than IOM needs for its purposes. Imagery and analysis should be based on the primary analysis given by the provider. Imagery in conflicts should only provide a general geographical overview of the delimited zone of interest.

To validate the images, they should be cross-referenced with other sources. For example, if the satellite images are embedded with geo-referenced information[19], they can be verified using Google Earth and other GIS / remote sensing software.

Conducting analyses

Limited access of imagery may be requested for primary analysis from the provider in order to extract data for analysis or fill out geographical gap of primary analysis. When analysing satellite images, do not assume the existence of an identifiable pattern; establishing patterns requires deep analysis from agencies and field reports. The direct analysis from satellite imagery should be used as a pattern and as basic information to support further analysis and field verification such as geolocation of IDP area settled, length of time, measurement of IDP area, building/shelter count and infrastructure level. The results of an analysis will need validation from IOM management and/or partners working directly with IOM in the area of interest.

Posting requests for imagery analysis in volunteer networks

Volunteer technical organizations are often used to analyse large amounts of information collected through satellite imagery. Volunteer networks should agree to keep data and analyses internal until a final agreement to share with the community. Staff should avoid posting requests for imagery pertaining to conflict areas in volunteer networks.

[19] Geo-referencing refers to assigning real-word latitude and longitude to each pixel of a raster.

Storage

Satellite imagery that is not public information should be safely stored and deleted when work is completed or per the data retention principles, as contained in the Migration Data Governance Policy (IN/253). A retention period must be set when making the request. Once the specified purpose is complete, the image should be deleted, unless there is a decision to extend the retention period for a specific time frame. Storage should be done per the level of classification. If the images received contain data that could be used to identify people, groups, or other confidential data, those things should be blurred from the image before storage. The storage of satellite imagery falls under the responsibility of the data manager. He or she will ensure that imagery is stored according to relevant guidelines and best practices.

Sharing

The sharing of satellite images should be based on a written agreement between IOM and any partners. The data steward is responsible and accountable for any such data that is shared internally or externally.

In order to share imagery not available to the public, an agreement must be signed with the provider of the images, as well as with the entity that the imagery will be shared with. Agreements for sharing satellite imagery should be coordinated with IOM HQ, as well as with LEG. If IOM is not the owner of the imagery, partners requesting for images should be directed to the initial provider.

In conflict contexts, updated imagery received for humanitarian purposes may include detail on displacement that is not available in the public domain. Even in non-conflict situations, it is advisable to thoroughly assess the potential sensitivities of the image when considering its use as background for a published document. It is advisably to make images used for backgrounds or for published documents very low resolution. Refer to the owner of the imagery.

Aim to prevent the misuse of satellite imagery by clearly indicating the purpose of the imagery along with the context, date, and which data is analysed. Any images should be marked explicitly whether they are for public or private use.

Unmanned aerial vehicles (UAVs)

The use of unmanned aerial vehicles (UAVs), commonly known as drones, in humanitarian contexts is rapidly growing. UAVs are usually equipped with tools such as cameras and microphones that can be used to gather visual data, audio, and thermal images. These data are often used, for example, post-disaster to identify damages to houses, roads, and other critical infrastructures, as well as identify the location of individuals who may still be at risk. The use of UAVs can greatly increase the speed and efficiency of gathering post-disaster data. For example, in 2012, serious flooding in Haiti, following Hurricane Sandy, left many areas inaccessible. In response, IOM Haiti deployed drones to conduct rapid damage assessments of the communities that were otherwise inaccessible. The aerial imagery provided by the drones, combined with existing open source imagery and census data[20], provided a complete analysis of the damage, just days after the event—7 days before satellite imagery was available. This allowed IOM to more quickly identify and respond to the specific needs of the affected population.[21]

The risks of using UAVs come, among others, from the technology with which UAVs are equipped. Some of the risks are, for example, the cameras, sensors, and microphones on UAVs can collect data that can be used to identify individuals or communities.

This section contains best practices in managing information collected, received or requested through the use of unmanned aerial vehicles (UAVs). It includes the use of UAVs in conflict areas, flight permission from authorities, information on flight times and data to be analysed, community and individual consent, real time imagery, tracking, high-resolution imagery and individual identity, invasiveness and deployment after disasters.

Design

Direct contact with the local community should be included early on in the design phase. IOM staff should work closely with the local community to explain and demonstrate the use of UAVs, dispelling any doubts or fears the community may have. The UAV mission should be used as an opportunity to produce information beneficial to the community, and planning for the flight should include consideration of what use will be invasive to the community.

IOM staff should avoid collecting a wide array of data for issues that are not under the scope of humanitarian action in general and under the scope of the specific project for which a drone was decided to be used.

IOM staff are advised to coordinate closely with other organizations requiring UAV data in order to avoid duplication and minimize the invasiveness of UAVs.

Flight permission from authorities

IOM staff should work with national and local authorities in order to ensure that the IOM UAV flight follows internationally accepted standards as well as any available local provisions. Given that UAVs are a new development in civil aviation, many countries have not yet developed relevant regulatory legislation.

All UAV flights should adhere to the ICAO RPAS Circular 328-AN/190 with the approval of national authorities. Flights without previous authorization run the risk of being prohibited.

In countries where there is no national regulation of UAVs, IOM staff should consider working with local authorities (e.g. the ministry or department responsible for civil aviation) and/or other organizations[22]

[20] Collected by IOM in partnership with the IHSI (Haitian Institute of Statistics and Informatics) for the Census of Affected Neighborhood and Population (RQPA) after the earthquake – to avoid confusion with previous census data from 2003.

[21] Case Study. *Mapping Rapid Damage Assessments of Tabarre and Surrounding Communities in Haiti following Hurricane Sandy.* Suisse Foundation for Mine Action (FSD). 24 June 2016. http://drones.fsd.ch/en/case-study-no-6-mapping-rapid-damage-assessments-of-tabarre-and-surrounding-communities-in-haiti-following-hurricane-sandy/

[22] Relevant organizations can include WFP (example http://werobotics.org/blog/2017/04/03/wfp-werobotics/), and the organizations mentioned in the FSD homepage (http://drones.fsd.ch/en/homepage/)

interested in using UAVs to develop instructions for using UAVs in the specific location and context. IOM can provide technical assistance as well as examples of existing regulations from other countries.

Local authorities should be informed of the flight times, data to be collected and analysed, and the organizations that will have access to the data. Transparent cooperation with authorities and local communities is paramount when using UAVs.

Collection

Community and individual consent

Community engagement and consent is particularly critical for longer term tracking of construction, camp closure or other activities that require multiple flights in the same location. The community should be informed of how information will be used and for what purpose. Local authorities should also be engaged on the use of the technology and the information collected.

Depending on the area where the UAV is flown, there might be individuals present in the area or not. If individuals are present, there arises issues of their consent to the UAV being flown. Consent is a major issue regarding UAV flights. Individual consent, in accordance with Principle 4 of IOM's Data Protection Principles might be difficult to achieve in an area where many individuals are present due to the nature of the data collection. Staff should provide information regarding flight times, data collected/analysed to the community members by signposting, leaflets etc, and provide a feedback channel through which communities can voice objections or concerns, or ask questions. Communication channels can be established through engagement with local authorities (including police), community leaders, and trusted community focal point and through local staff.

Staff should consider the security context and in cases where staff cannot visit communities directly, should disseminate information regarding flights and data to be collected and analysed through local authorities, whether national, regional or through other organizations.

Use of UAVs in conflict situations

UAVs will not be flown by IOM in areas affected by conflict. The reasons for this principle are complex, but are largely concerned with the duty of care towards IOM beneficiaries and IOM's neutrality at national and international levels.

Real time imagery

Real time video imagery from UAVs is available, but should not be used unless there is a specific program objective that requires and justifies its use. Real time imagery can be difficult to obtain and cause unnecessary risks for the pilot of the drone or the affected population. Since drones are not usually flown under long range control, piloting them close to an affected area can be dangerous.

Technical specifications

Recording settings such as the resolution of imagery captured can be modified in order to protect privacy. Methods used to protect privacy may vary according to the goals of the specific project, whether, for example, the aim of the data collection is to survey an area, a structure or to match GIS datasets, each requiring a different level of definition in images. UAVs can be flown in nadir (vertical positions), low oblique, high oblique or LiDAR (360 degree mode used to capture 3D imagery). The best mode to use for mapping is nadir or high oblique as it provides the best information, while avoiding compromising individual privacy. LiDAR is a laser sensor that is costly and needs bigger drones or even airplanes and therefore is not available for IOM's use. For 3D imagery, IOM staff can process overlapping nadir or oblique imagery and use software such as Pix4D to reconstruct a 3D model[23].

Analysis

Cleaning and validation

While using angles to protect individual identity, the UAV can still pick up specific information that can be considered invasive. In this situation, the data in question should be cleaned, removing any data considered to be invasive-based on who will be using and viewing the data.

In order to define what can be considered invasive (e.g. property lines), IOM staff are advised to conduct a field appraisal or analysis of the area in which the UAV flight is executed and to identify the socio-cultural context of the area, as outlined in the design phase above.

Conducting analyses

Before commencing the analysis, the objectives and variables for the analysis should be discussed and agreed with the involvement of the local authorities. This prior definition will ensure that the scope of the analysis is limited to the required information and its use in project implementation. Analyses should include baseline data that is normally obtained through satellite imagery in order to accurately identify changes.

When analysing geo-spatial data to get up to date information of affected areas, staff should employ the BAR methodology, as promoted by the Signal Program. This standardized process includes: setting the parameters, assigning structure categories, assigning damage scale, and calculating point totals. Staff should consult the *Imagery Interpretation Guide: Assessing wind disaster damage to structures*[24] for complete guidance on each of these steps. [25]

Storage

Storage and management of UAV imagery

UAVs create a large amount of visual data that may include sensitive information. The question of long-term storage, archiving or deletion of such footage should be carefully taking into consideration the how the data is classified. Storage of such data should also include the prioritization of community needs.

[23] See https://sketchfab.com/models/df2663dea2dd42fc8fd8872223f3c57c for an example

[24] Achkar. Baker. Raymond. *Assessing wind damage to structures.* Harvard Humanitarian Initiative. http://hhi.harvard.edu/sites/default/files/publications/siig_3_22mar2016.pdf

[25] Examples of IOM UAV activities as a resource for analysis: https://onedrive.live.com/?authkey=%21AOQE2F9OtNFVtBI&cid=7C6901C14D035FEA&id=7C6901C14D035FEA%2134930&parId=7C6901C14D035FEA%21112&o=OneUp; https://onedrive.live.com/?authkey=%21AK5FcpQwftApIWM&cid=7C6901C14D035FEA&id=7C6901C14D035FEA%2122123&parId=7C6901C14D035FEA%21112&o=OneUp; https://drive.google.com/file/d/0BwfYtgMJgxLaQXVKRnEwaXhsNW8/view

Personal data collected via UAVs should not be stored longer than is necessary to fulfill the specified purpose of collection, in accordance with IOM Data Protection Principles. Once the purpose for which it was collected, it should be deleted or anonymized.

Storage of hardware

IOM offices must establish usage, storage and security protocols, according to exciting internal policies, to avoid loss, theft and misuse of costly UAV hardware.

Sharing

Given its potential sensitivity, UAV data should only be shared if the context permits. If this is the case, sharing should only be to trusted partners with adequate cyber security and data responsibility mechanisms in place, and for a pre-specified purpose, in accordance with a written agreement or memorandum of understanding. The terms and conditions for sharing UAV data should cover at least: means of transfer, storage, access restrictions, any publications, and the deletion of the data.

If shared, this imagery should include a disclaimer with the logo, the source organization and the date. The level of detail may be decreased before sharing. Before sharing, the initial purpose should be established as the detail needed may vary depending on the purpose of the imagery, i.e. the level of detail needed to be used in GIS software is different than if producing a document or a PDF.

Crowdsourcing

Crowdsourcing refers to “the act of taking a job once performed by [staff], and outsourcing it in the form of an open call, to a large group of people”[26]. Staff can incorporate crowdsourcing techniques and technologies to facilitate different phases of the data processing cycle. In this guidance document, crowdsourcing is divided into two main functions:

1. User-generated information

Beginning with data collection, staff may use several different channels such as social media, SMS, and instant messaging applications to communicate with affected populations, both to disseminate and to receive information. One practitioner explains crowdsourcing as an upside-down umbrella in a rain storm. As rain falls, it is captured through the extended spokes of the umbrella and funneled to the center.[27] The information provided through crowdsourcing channels is referred to in this document as *user-generated information*[28]. This document identifies two ways in which user-generated information is gathered. First, it can be gathered through targeted information activities, during which users voluntarily provide information to humanitarian actors through the specified channel. Second, user-generated information

[26] Howe, Jeff. *The Rise of Crowdsourcing.* Wired Magazine. January 2016. Accessed at: https://www.wired.com/2006/06/crowds/

[27] Rotich, Juliana. *Crowdsourcing for Humanity* (Interview with ICRC). Published September 2013 at: https://www.youtube.com/watch?v=iLh4kB3doLQ

[28] *User* may not be limited to the affected population only.

can be obtained through analyses of datasets, either publicly available or made available to humanitarian organizations by service providers such as social media companies or network operators.

2. Volunteer networks

Immediately following crises, there is usually a huge influx of user-generated information relating to the incident. It may be impossible for the organization to process such vast amounts of data. The use of volunteer networks allows staff to outsource specific technical tasks such as validation, analysis and visualization of data to a crowd or network of volunteers that may be situated anywhere in the world.[29]

This guidance document will begin with a general overview of the uses and challenges of user-generated information, as well as general recommendations to be applied when using user-generated information. In the following sections, guidance is given on specific technologies and channels that are used in sourcing information from a population, giving information on the uses, challenges, and best practices of social media, instant messaging applications, and SMS communications. Finally, this document presents considerations to be taken when working with volunteer networks to perform specific technical tasks.

Activities to crowdsource information from affected populations, as well as with all external communications, including communications with communities should always be coordinated with MCD.

Challenges

Processing

As mentioned above, the main challenge of crowdsourcing and user-generated information is the amount of information it will produce. Crowdsourcing activities can result in a flooding of information, not all of which is relevant or accurate. Processing such diverse, large amounts of data will require a significant investment and may lie beyond the capacity of the project.

Bias

Information collected through crowdsourcing can be heavily biased to populations that are adept and have access to technologies. Basing assumptions of the entire affected population based upon information gathered through social media or other channels can be problematic in that some the population, often the most poor and vulnerable can remain completely invisible.

Safety

Crowdsourced information may have implications for the safety of the respondent. This is especially the case in instances of violence as either repressive governments or other groups could potentially use the information to identify a vulnerable individual, community, or demographic.

[29] UNOCHA. Humanitarianism in a Network Age. Pg 29

General recommendations

Design

Identify the exact purpose of the data to be collected in order to help avoid data overload or irrelevant information. The potential respondents need to be informed on this purpose so that they only provide data that will serve a specified purpose.

Understand how information flows among the affected population. Identify any instances of differences of communication channels based on access to technology, age, sex, socio-economic status, and geographic location that could present biases to crowdsourced data. Diversify the channels used accordingly, using mobile, online, and offline mechanisms to gather information. This will help mitigate the risk of getting skewed data based on access to or use of technology.

The information collected, especially through those opting in to information campaigns, should be limited to the project's capacity to process and respond to the information collected.

Collection

Consent

A well-defined collection process will help to ensure the validity of the data and the safety of the respondents.

In calls for user-generated information, specify the purpose for which the data will be used. This will help manage the expectations of the users.

Information can either be generated voluntarily as a user responds to a specific call for information, or unknowingly if the data is gathered through publicly available social media posts, for example, or from datasets obtained from a private company. For those voluntarily providing information, total transparency is needed for the risks associated with providing the data, as well as how the data will be used and processed. Respondents should understand that in cases where providing information can be a threat to their safety, it should be avoided.

Data supplied by third-parties

Anytime data is supplied by third parties, including through crowdsourcing methods, it is essential to verify that the organization or group supplying the data has used ethical practices in their digital research. In order for data or findings to be accepted for IOM projects, the supplying organization should be able to outline and describe the ethical methods that were used to collect the data.

Analysis

Cross referencing systems should be in place to rapidly identify inaccurate information, whether intentional, exaggerated, or accidental. Information that is crowdsourced can come from many different media, some easier to verify than others.

The quality of information can be monitored by mapping information to see if there are any differences based on the source of the information Measures should be in place to remove or edit any uploaded data containing false or identifying information.

Systems should be in place to triage and prioritize large amounts of information. This can mean identifying key words and phrases to incite attention to a specific item.

In instances where staff do not have the capacity to analyse the data, it may be recommended to conduct advanced computing to verify user-generated content. Advanced computing comes in two forms: 1) human computing, and 2) machine computing.

Applying advanced computing to verify user-generated content

Human computing

Human computing refers to the process of micro-tasking specific aspects of data verification volunteers.[30] This can be used to quickly perform tasks such as the verification user-generated content or sorting which information is relevant and non-relevant to prioritize what is needed for further analysis. More information on using volunteer networks for human computing is given in the *volunteer networks* section of this document.

Machine computing

Machine computing includes software and algorithms designed to rapidly verify large amounts of information. Machines take into account factors such as length of posts, punctuation and language usage in order to predict the accuracy of user-generated content.[31] Machine computing can greatly reduce the time needed to analyse data. Further guidance is given in the section of this document dedicated to advanced computing methods.

Storage

Crowdsourced data should not include personal data. If they do, IOM must immediately delete such personal data. Guidance for the storage of data collected through specific channels is given in the subsequent sections of this document.

Sharing and use

As previously stated, crowdsourced information is usually available to a wide audience, so it should be clear in the design phase how the data will be shared, used and accessed.

Along with sharing information with relevant partners, mechanisms should also be developed to communicate information to the affected population, eliminating the one-way flow of information common in crowdsourcing.

Social media

People affected by a crisis or disaster often turn to social media platforms such as Facebook, Twitter, Instagram, to give information and communicate their needs. This has been attributed to three reasons.

[30] Collins, Kate. *How AI, Twitter and digital volunteers are transforming humanitarian disaster response.* Wired. 30 September 2013. http://www.wired.co.uk/article/digital-humanitarianism

[31] Meier, Patrick. *Adding the Computer Crowd to the Human Crowd*. Verification Handbook (chapter 7). January 2014

First, social media has the potential to reach a wide audience. This is especially relevant if there is limited band-with or power following a disaster. Second, social media posts are likely to stimulate interaction. Those who post on social media usually expect comments and engagement from those within their network. Finally, social media is often seen as increasing the likelihood that user-input will have an impact.[32] This provides a unique opportunity for humanitarian organizations to leverage social media for a more informed response, giving access to real-time information that is generated by the affected population themselves.

One example of using social media to improve disaster-response is the IOM-developed *Community Response Map* (CRM). CRM is a platform that allows staff to communicate directly to populations through social media, public events, and other social marketing tools. The affected population can provide feedback, through user-generated information, that can trigger an urgent response, or a follow-up action. CRM turns simple information campaigns into an interactive feedback mechanism.[33]

Another use of social media by IOM staff was given in the survey. In the example, it was suspected that the total number of people affected by a particular crisis had been inflated and did not represent the actual situation. The IOM office was able to work with the national government and analyse social media data to provide a more accurate number that better represented the actual people affected.[34]

Challenges

Verification

In the influx of user-generated information immediately following a crisis, verifying the content of a social media posts can be particularly challenging. Significant time investment is needed to analyse crowd-sourced data through social media. Social media content is user-generated, meaning anyone can post what they want, regardless of whether or not it is factual. Content and photos may be altered intentionally, to increase the popularity of a user's profile, or unintentionally, as users share or re-post false information thought to be correct.

Bias

Another challenge is the differences in access to social media. While social media is becoming increasingly popular, there is still a significant number of the world's population that do not have access to it. IOM staff should remain cognizant of the potential for bias due to differential levels of social media penetration by geography, infrastructure, socio-economic group, or demographic. Finally, using a social media platform can pose threats to the privacy and security of the user as posts are usually public, providing information to individuals or groups who may want to cause further harm.

[32] Ward, Amy. *Social Media in Disaster Response*. Citizen Tech. http://amysampleward.org/2011/02/19/citizen-tech-social-media-in-disaster-response/

[33] https://communityresponsemap.org/how-it-works

[34] Responses from the staff survey are to remain anonymous

IOM Guidelines for the Use of Social Media

When using social media for a project, IOM staff should adhere to the Guidelines for the Use of Social Media[35] (Social Media Guidelines), a routinely updated document available to IOM staff. The Social Media Guidelines provide information on ensuring privacy and security, managing inaccurate information, and best practices when using Facebook, Twitter, YouTube, and Instagram. Each staff is responsible for understanding and applying the contents of the Social Media Guidelines when using an IOM social media account.

Provided below are best practices for managing social media information throughout the data processing cycle.

Design

Before beginning any activities to collect social media information, an analysis should be done to establish both the utility of the information to be gathered, and whether or not the IOM office has the adequate resources and capacity to deal with the vast amounts of information that can be generated through social media. Not only will the amount of information be extensive, but it will expand across many different types of media, such as text, video, and images. If this information will not feed into a targeted response, or if the IOM office does not have the resources necessary to handle such a vast amount of information, it is best not to proceed.

Standardization of hashtags

Official hashtags should be identified and promoted as early as possible, in coordination with other humanitarian organizations and the government. This will enable better data aggregation and analysis, reducing the amount of information to process, creating a timeline of events and tracking needs, people and events, as well as supporting evidence of users providing their consent. When standardizing hashtags in disasters, three types should be used, a disaster name hashtag (e.g. #haiyan), a public reporting hashtag (e.g. #PublicRep) and an emergency response hashtag (e.g. #911US).[36] However, staff need to remain flexible and adapt to the dynamic social media context. Not everyone in the target population will use the promoted hashtags. Some may develop their own, while others may customize those proposed to better reflect the situation.

Geo-targeting

Geo-targeting, or tailoring communications to a geographical area, can be used to specifically target information to individuals who need it the most. Geo-targeted hashtags and/or content may help to increase the relevance of the message and reduce the likelihood the message will be ignored. It also helps avoid over-alerting, or sending too many messages to those who may not be affected by an event. Furthermore, if everyone receives the message, it is less likely that those actually at risk will ignore the message, as they assume a generic message may not apply to them. This will help both the affected populations to report their needs, as well as the humanitarian community in coordinating a response.

[35] Available to IOM staff through IOM intranet: https://intranetportal/Pages/ControlNo.aspx?controlNo=SD/MCD/00047

[36] UNOCHA, Hashtag Standards for Emergencies. 2014.https://www.unocha.org/sites/unocha/files/Hashtag%20Standards%20for%20Emergencies.pdf

Word analysis

Before starting, IOM staff should map out all necessary keywords and websites targeted, determine the sample size and the humanitarian purpose for which the collection and analysis will be used. Staff are advised, in order to ensure ethical data analysis, to remove personal information that can identify individuals or groups and increase their vulnerability.

Collection

Consent in social media

Emergencies and crises often lead to a deluge of information regarding needs of the affected population. This voluntary provision of information on social media networks could be considered a voluntary opt-in action and therefore implicit consent provided by the person providing the information in the public domain.

The argument in favor of implicit consent is strengthened by the fact that social media users provide the information on the public domain and they have agreed to the terms and conditions of the specific social media platform. It could also further be strengthened if the individual providing the information uses the standardized hashtags popularized by IOM social media accounts.

However, IOM staff are advised to contextualize user generated data via social media to gauge whether there are reasonable indicators that the social media user is opting into the data collection process. Key concerns include whether any other third parties are identifiable as well as political and security implications. Consent of the social media users for allowing IOM to collect and analyse their data is a complex issue. In case of doubt, IOM staff should consult with the policies and terms of use of the specific social media platform and discuss with LEG.

If possible, staff may pre-identify a group of trusted informants who, with their consent, are able to provide legitimate information within a specified timeframe and with a specified regularity.

Analysis

Verifying social media information

The ability of anyone to upload content to social media makes it very difficult to establish what is authentic and what is not. IOM staff should ensure that data gathered from social media is vetted for accuracy and relevance. False user accounts, hoax websites, Photo editing software, and fake news further complicate the verification process. While some take extreme measures to deliberately supply false information, some are simply careless and do not realize themselves that the content is false. It is therefore imperative to verify information gathered through social media. Before using information from social media, staff should be able to verify the following aspects of the provided information[37]:

[37] Silverman, Craig et al. *Verification Handbook: Verifying Digital Content for Emergency Coverage.* (2013) http://verificationhandbook.com/book/chapter3.php

Provenance	•Is the content original? (not "re-posted") •In case of images, try reverse images seaches such, TinEye, RevEye or Google images
Source	•Who supplied the content?
Date	•When was the content uploaded? •Use tools such as WolframAlpha and SunCalc to verify local weather conditions
Location	•Where was the content uploaded? •Use tools such as Google Maps, Google Earth, Open Street Map, WikiMapia and GeoNames
Validity	•Is the content consistent with other information? •Triangulate the information

Crisis computing

Given the mass amounts of information that flood the internet with the occurrence of natural disasters or emergencies, cyber disinformation can be a threat. In this scenario, and in order to determine the credibility of this data, IOM staff may employ crisis computing. In this case, the veracity of information should be tested either by crowdsourcing (through digital volunteers) or through algorithms (using a significant quantity of data with certain degree of accuracy). More information can be found in the advanced computing section of this document.

Conducting analyses

Aggregators and algorithms for analysis

Social media networks provide a large array of scattered data. By using networks of volunteers and other advanced computing mechanisms, the data can become localized and searchable. In order to ensure that the dataset is not utilized for other purposes and to avoid "functional creep", IOM staff are advised to ensure the anonymization and disaggregation of data as far as possible and as soon as possible in the data processing cycle. As analyses can present many challenges, IOM staff are advised to quickly determine and only utilize the data necessary in order to meet the needs of the specific project.

Personal information, such as address, date of birth, and identifiable photographs should be removed before conducting analyses.

Storage

Despite the public nature of user-generated social media information, it should still be handled with care. Information that has been gathered through social media should be stored on a secure server and classified in accordance with the IOM Data Governance Policy (IN/253).

Sharing and use

Source and privacy of user-generated data

Some user-generated data may be sensitive[38]. Staff should be mindful to not disseminate sensitive data which would prove to be a risk for any individual or group. Data may relate to the user themselves, or to third party vulnerable individuals or groups. Prior to its dissemination, data should be anonymized to mitigate risks.

It is advisable not to re-publish social media data in raw form, or to re-publish personal profiles, even if already public. Only anonymous statistics should be published.

Mobile Networks

Mobile networks can provide humanitarian data through two main functions: Short message services (SMS), and mobile network data such as call detail records (CDR).

Short Message Services (SMS)

SMS is a communication service available on most mobile telephony systems.[39] It enables users to communicate in short, mainly text-based messages.[40] Humanitarian action often includes SMS communications to allow direct communication with an affected population, providing the necessary channels to communicate locations, and urgent needs.

Mobile network data

Mobile network data refers to the anonymized information provided by Mobile Network Operators (MNOs) on users' SIM cards. The data collected is referred to as call detail records (CDRs), showing the times and locations where SIM cards were used to make calls, send or receive text messages, or download data. This information is especially useful to track population movements. CDRs for example can determine movement by tracking where calls were placed from a particular SIM card before and after an event.

Benefits

With so many users of mobile phones, SMS and mobile network data provide many uses in humanitarian operations. SMS can be used to supply timely information, as well as gather information to fuel a response. Following crises, many people turn to their cell phones to contact loved ones and communicate their needs. Mobile phones the most common means of communication among affected populations, [41] making CDRs particularly useful in humanitarian operations.

[38] Whether personal (names, social security numbers etc.) or non-personal (GBV cases, number of UASC in a region) etc.) in nature.

[39] Operation and use of mobile telephones

[40] https://en.wikipedia.org/wiki/SMS

[41] GSMA, Towards a Code of Conduct: Guidelines for the Use of SMS in Natural Disasters.

To give a famous example, following the earthquake in Haiti in 2010, Usahidi, a software company that develops free and publicly available software for information,[42] partnered with the Haitian telecommunications company *Telcos,* and the Emergency Information Service of Haiti to conduct an SMS information project. In the project, affected individuals simply had to text their location and urgent needs to the "4636". The text message was sent, at no cost to the user, to Ushahidi. The information was then translated, mapped, and prioritized for the response; all within minutes of the message being sent.[43] CDRs have the potential to provide near-real time information on population movements leading up to, or following a disaster. One example comes from Nepal in the aftermath of the earthquake in 2015. Call detail records from 12 million de-identified mobile phone users were analysed and able to help identify the movements of the affected population. CDRs were able to give an unprecedented level of information in an extremely short period of time that would not have been available through any other means.[44]

Challenges

Spam

Those who receive the messages may consider them as spam and ignore them. This especially true if there are multiple organizations engaging in SMS activities, resulting in populations receiving bombardments of similar messages.[45]

Bias

As with other technologies, not everyone in an affected area may have access to a telephone, potentially skewing the data to a particular demographic. CDRs may also be skewed in the instances where SIM cards are quickly replaced, when multiple people use the same SIM card, or when one person uses multiple SIM cards.

Power outages

Following disasters, power-grids may be damaged, leaving mobile phones un-useable once they have lost their charge.

Network issues

SMS requires cellular network that may not always be available in the affected area. One reason for this, is the damage that a disaster may have on the infrastructure, including the cell towers. In this case, the MNO will likely prioritize restoring their own infrastructure before aiding in humanitarian operations. In addition, in rural areas, or places with low infrastructure development, there may not be adequate cell network. Finally, immediately following a disaster, too many users can overload the network if they try to access it at the same time, making calls and texts impossible.

[42] https://www.ushahidi.com/

[43] Meier, Patrick. *Ushahidi & the Unprecedented Role of SMS in Disaster Response.* iRevolutions. 20 February 2010 https://irevolutions.org/2010/02/20/sms-disaster-response/

[44] Wilson, Roberts et al. *Rapid and Near Real-Time Assessments of Population Displacement Using Mobile Phone Data Following Disasters: The 2015 Nepal Earthquake.* PLOS. February 24, 2016. http://currents.plos.org/disasters/article/rapid-and-near-real-time-assessments-of-population-displacement-using-mobile-phone-data-following-disasters-the-2015-nepal-earthquake/

[45] IFRC. *Two-Way SMS Communication with Disaster Affected People in Haiti* (case study). March 2012. http://reliefweb.int/sites/reliefweb.int/files/resources/trilogy_intl__ifrc_case_study.pdf

Design

Begin by identifying if an SMS activity is appropriate for the specific context. Identify differences in mobile phone ownership, access to network, literacy levels, and other factors that can affect the usage of mobile phones. Understand the gaps and biases that may render a certain demographic invisible. Setting up any mobile network data or SMS communication system should take place as part of preparedness activities, making sure systems are in place before a disaster strikes.

Coordinate

Communicate plans to use SMS communications or CDRs to relevant organizations and government entities. If there are any existing government SMS alert programs, work within the existing systems rather than duplicating them. Coordinated approaches to MNOs together with other humanitarian actors will help avoid sending multiple requests and increase the likelihood of partnership.[46]

Ensure Capacity

Verify that the office or mission has the capacity to respond to issues raised in SMS activities. Information should only be gathered on items that IOM has the capacity to respond to, otherwise, false expectations will be raised and have a negative effect on the relationship with the population as well as the reputation of IOM. Relatedly, IOM offices should assess whether they have the necessary capacity to process the potentially large amount of information gathered through SMS-communication initiatives.

Partnership with Mobile Network Operators

In many instances, SMS activities will require establishing partnerships with Mobile Network Operators (MNOs). Not only do MNOs own the network, they can also provide the expertise in delivering effective messages to their customers.[47] It is important to keep some key considerations in mind when establishing partnerships with MNOs.

Partnerships should be set up ahead of time. After the onset of a disaster, it is normally too late. Partnerships should include formal written agreements with MNOs, establishing the timing and level of detail to be shared in data sets. All agreements should be coordinated with LEG.

The company itself may be badly affected after a disaster; it should be assumed that the company will give priority at this time to restoring their own services.

Identify a single point of contact between IOM and the MNO.

Consider how the intended program can fit into pre-existing programs the MNO may have in place with local NGOs. Partnerships with MNOs should be established well in advance of a disaster and included in preparedness activities.

[46] ibid

[47] *Towards a Code of Conduct: Guidelines for the use of SMS in natural disasters.* GSMA. https://www.gsma.com/mobilefordevelopment/wp-content/uploads/2013/02/Towards-a-Code-of-Conduct-SMS-Guidelines.pdf

Crowd seeding

Crowd seeding[48] can be used to overcome differences in access to SMS communications and help ensure the validity of the gathered information. This refers to pre-selecting key informants from among a population, training them to provide accurate, representative information and providing them with mobile phones and adequate airtime to communicate relevant updates. When possible, key informants should be trained ahead of time as part of preparedness activities.

Collection

SMS

Consent

A legitimate purpose of the information to be collected should be communicated to the users at the beginning of the project. Those who participate should understand that providing information implies their consent for that information to be used for the specified purpose.

Simplicity

Emphasis should be placed on simplifying the messages and minimizing the information received. Specify the type of information that is needed for the purpose of the project (e.g. location and need). This will decrease amount of information provided, as well as the needed time investment to analyse the information generated. As mentioned above, only seek to gather information that falls within IOM's capacity to respond, so as to not raise false expectations.

Feedback

The service should be coupled with complaint mechanisms for users to provide their feedback. In addition, include an opt-out option on the messages to allow the users to unsubscribe at any time from specific messages.

Geo targeting

Geo targeting messages, as with social media, to only affected areas, will help decrease the resources needed to process the information, as well as reduce message fatigue for those who do not have information to contribute. Geotargeting can be achieved by identifying which cell tower a user is connected to, messages should only be sent to those in specific areas.

Mobile network data

Staff should not acquire raw CDR data from MNOs. MNOs are usually bound by strict protection and privacy regulations that prevent them from sharing personal information, such as the user's phone number. The MNOs will always store the raw data on their own servers and only allow analyses to take place under their supervision and only on de-identified and aggregated data.

[48] Vanderwindt, Peter. *Crowdseeding in Eastern Congo: Using Cell Phones to Collect Conflict Events Data in Real Time.* Journal of Conflict Resolution. 2014

Analysis

Though the messages may be coordinated through an MNO, the responsibility to verify the information falls on IOM. Information should always be triangulated and tested against others sources for validity.

Personal information should be removed from the message content.

SMS software packages

There are various software packages designed to facilitate mass-SMS communications. Programs such as FrontlineSMS[49] connect standard cell phone to a laptop, coordinating the sending and receiving of vast numbers of messages. Once messages are received they are analysed and coded based on location and content. The software is able to create databases and visualization charts based on the data. This can greatly reduce the time and capacity needed to process the data. However, using the software does not eliminate the need to verify and validate the information received.

Preprocessing (CDR)

In order to avoid an unmanageable deluge of information, some steps can be taken to reduce the amount of information to process. This can be done by including preprocessing algorithms in analysis tools. These programs will remove all unused information, significantly reducing the information to a more manageable size.

Recommended preprocessing activities

Remove unused information. Only include the time of phone call or message and the location in analyses.

Assign a cell tower a location ID and aggregate based on administrative boundary levels. For example, if one boundary has 20 cell towers, all 20 towers will be represented by one location ID in the dataset.

Define how the daily location will be identified. Rather than analysing each separate activity from a SIM card. Choose a time throughout the day such as, "the last call made in the day" to define the location for that day.[50]

Analysing CDRs

Analyses are made by remotely connecting to an MNO's server and is limited to aggregated, de-identified data. Any personally identifiable data should never leave the MNO's servers and should not be included in the analysis. Data made accessible to IOM should only include the times of activity and low-resolution location data (to the nearest cell tower).[51] Only the aggregated data that provides the estimates of mobility should be exported off of the MNO's servers.

[49] http://www.frontlinesms.com/

[50] Wilson, Robin et al. *Rapid and Near Real-Time Assessments of Population Displacement using Mobile Phone Data Following Disasters. The 2015 Nepal Earthquake.* PLOS. February 24, 2016. http://currents.plos.org/disasters/article/rapid-and-near-real-time-assessments-of-population-displacement-using-mobile-phone-data-following-disasters-the-2015-nepal-earthquake/

[51] Bengtsson, Linus. *Mobile Operator Call Records: Potential and Pitfalls.* Flowminder.org. https://www.oecd.org/sti/ieconomy/1%20-%20Linus%20Bengtsson.pdf

Not all of the movement after a disaster can be attributed to the disaster itself. Use the data to establish the baseline population flows pre-disaster and compare to the flows after the event.

Storage

Mobile operators reserve the right to store CDR on their own servers, allowing only those with specific and limited access to conduct any analysis. Only the results of the analyses will be allowed to be migrated off their servers. Where access is granted to such data, and allowed to be moved to other servers, it must also be stored on a secure server.

In the case of using SMS communications, message content should be classified as confidential and is subject to the specifications for retention and disposal as outlined in IN/5 and IN/123.

Sharing

Only the results of the analysis, such as aggregated statistics, or non-sensitive data on mobility estimates should be shared with other actors, agencies or governments.[52] Any sensitive data or potentially identifiable information from CDRs or from SMSs should not be shared.

Instant messaging applications

Instant messaging applications (IMAs), such as WhatsApp, Viber, Snapchat, WeChat, Kakao and Facebook Messenger are "mobile phone-based software program[s] that allow users to send and receive information using their phones"[53]. In comparison to SMS, messaging apps facilitate sharing many forms of media, including sound files, texts, videos and images. The popularity of instant messaging applications (IMAs) has risen drastically in recent years with the increase of accessibility of smart phones due to reduced prices of mobile data and smart phones. This has drawn many humanitarian organizations to use messaging apps to better communicate with affected populations, and to coordinate tasks internally.[54]

Benefits

Low cost

Using messaging apps can reduce communication costs within the organization as it is usually far cheaper to send a message through an internet connection as opposed to messages sent through mobile networks. Instant messaging apps also facilitate remote communication in cross-border response situations, as it costs no extra money to cross a border or work with a different network, as is the case with SMS communications.

[52] GSMA-Guidelines on the protection of privacy in the use of mobile data for responding to the Ebola outbreak. 2014 https://www.gsma.com/mobilefordevelopment/wp-content/uploads/2014/11/GSMA-Guidelines-on-protecting-privacy-in-the-use-of-mobile-phone-data-for-responding-to-the-Ebola-outbreak-_October-2014.pdf

[53] Foran, Rose. "*The refugee crisis will not be hacked*", 29 July 2016

[54] Lindsey-Curtet, Charlotte. *Messaging apps: an untapped humanitarian resource.* (News release from the ICRC) 31 January 2017. https://www.icrc.org/en/document/messaging-apps-untapped-humanitarian-resource

Connectivity

In many cases, following disasters when other methods of communication have been destroyed or left otherwise non-operational, internet connections have remained in place, allowing for continued communication through messaging apps.[55]

Security

Messaging apps can potentially be more secure than traditional SMS or phone calls, as some IMAs are able to encrypt conversations (see section on encryption below for more information).

Challenges

Information overload

Messaging applications can provide channels to share data in many formats, text, video and audio files. Humanitarian actors implementing communications through messaging apps may find themselves overwhelmed with the volume and diversity of the data to be processed.

Bias

It is true that messaging apps are rising in popularity, however, they are only available to people who own smart phones. In addition, IMAs require a WIFI connection, or a data plan. This can potentially exclude a major part of the population. Differences in age, socioeconomic status, gender, and technological infrastructure all play a role in the accessibility, affordability, and usage of smart phones.

Lack of knowledge

In general, there is little knowledge on the use of smart phones, messaging applications, and the overall technological ecosystem in crises, as access and use of technology are seldom included in needs assessments.

Privacy

Data that is collected through messaging apps is subject to the policies and standards of the company that owns the app. Companies vary in the extent of the information of their users they store on their servers. Data can vary from the location and time of messages, to the actual content of the messages themselves. These data can potentially be used by third parties to identify individuals or groups.

Design

When considering the use of messaging apps for communicating with communities, staff need to understand the technological ecosystem of the area, identify the appropriate messaging application, and know how to best mitigate the potential risks to the privacy, data protection, and security of the invividuals providing information through the app. As with any technology, the staff should ensure there is adequate capacity required to process and take action from the data.

Understand the ecosystem

It is essential to understand the population's access to and usage of technology. Indicators to identify the affected population's preferred methods of communication should be included into needs assessments.

[55] *Humanitarian Futures for Messaging Apps.* ICRC, the Engine Room and Block Party. January 2017.

Differences in usage of feature phones and smartphones should be measured across different demographics, allowing for the compensation for gaps in access to technology.

As with all other methods of communication identified in this document, messaging apps should never be the sole means of communication or gathering information, and their use should only make up a portion of the methods or technologies used to gather information. It is recommended to carefully consider the socio-cultural context in which the data collection takes place to judge which app provides a safer and more productive communication with key informants.

Choosing an app[56]

When possible, choose a messaging app company with policies and standards that are most likely to ensure the privacy and security of its users. This practice can sometime be a balancing act; identifying messaging apps that are familiar to the affected population, while ensuring the app used fulfills all of the security criteria. Some key security features of messaging apps to consider are as follows:

Data retention

The company should maintain a standard of no or minimal retention of information on their servers. Usually, the less data about the users the messaging app stores on their servers, the safer they are as an option. Companies that store the contents of messages should be avoided in particular. WhatsApp, Viber and Signal delete the content of messages from their server as soon as they are read by the recipient. However, Skype, and Facebook Messenger keep all message content stored indefinitely.

End-to-end encryption

Choose a messaging app that uses end-to-end encryption, this prevents any third parties, including the app-company, from accessing message contents. Messaging apps with end-to-end encryption as a default setting are preferred, this includes: WhatsApp, Viber, Signal Telegram, Line and Firechat. End-to-end encryption is also available on Facebook Messenger, but must be enabled using the *secret conversation* option. Snapchat, Skype, imo, and WeChat offer no end-to-end encryptions and should be avoided.

Data sharing

Choose a messaging app that does not share data beyond that which is minimally necessary to third party providers. Understand that apps do need to share some data to third parties who provide technical services, however, apps that share personal data of their users should be avoided. The complete list of messaging apps, as well as how they measure against all the considerations, are available in the publication *Humanitarian Futures for Messaging Apps.*[57]

Application development

Application development initiatives for functionality not yet served by corporate applications or for project related purposes, must be coordinated with ICT for verification of design against security, software

[56] Expanded guidance given in: *Humanitarian Futures for Messaging Apps.* ICRC, The Engine Room and Block Party, January 2017 (pg. 67-69)

[57] https://www.icrc.org/en/publication/humanitarian-futures-messaging-apps

development standards, existing applications for compatibility with other IOM systems. Full details can be found in the ICT Software Development Standards document on the ICT intranet site.

Collection

Consent

Inform and ensure consent of potential users from the outset regarding all the intended uses of the data. Remember to only include the minimum information needed for the intended purpose. User should be informed as well on their option to remove their information, although this can be difficult if it is stored on the company's server.

Define the content

To limit the amount of information, especially personal information shared through a messaging app, state explicitly the type of information that is needed and what it will be used for.

Unsubscribe option

Include in any communication, the option to unsubscribe, or opt-out of further communications.

Analysis

Verifying the data received through messaging apps can be complicated due to the speed at which information is received, the volume of information and the variety of data types that can be sent. Verifying information through messaging apps should follow similar steps of verification for other user generated content as outlined above in the social media section of this document.

Note that many systems used by messaging apps are often not interoperable with existing information management systems or databases. Therefore, steps should be taken to ensure that data is received in a uniform format, reducing the investment of time needed to sort through different types of data.

Storage

Many IMA companies retain information from their apps on their own servers. It is important to choose an app that deletes the contents of a message from their servers once the reader as accessed the message. The content of messages should be considered confidential and handled in accordance with the IOM Migration Data Governance Policy (IN/253).

Conversations through messaging apps may also be saved automatically in the chat history on the mobile device itself. Governments and border authorities may be suspicious of smart phones and messaging apps. There have even been instances of authorities demanding the passwords of electronic devices[58] of humanitarian workers. Staff should adopt the practice of regularly clearing chat histories on devices used to communicate with affected populations that may contain sensitive information.

[58] Al Jazeera http://www.aljazeera.com/indepth/opinion/2017/05/prison-privacy-170517123142625.html

Sharing

Only the analyses of messages should be shared with partners. Any data gathered using messaging apps should be de-identified.

Consider that many messaging app companies reserve the right to share some information that is necessary to third parties to perform specialized tasks, so it is best to choose an app that shares minimal information.

Advanced computing

As the use of technologies has increased, so has the amount and variety of data available to companies, governments and individuals. Each time an individual interacts with a technology, makes a transaction, uses social media, connects to the internet, and countless other things, data is generated and stored. The vast amount of data that is passively collected in similar ways is referred to as big data. Big data is characterized using the "4 V's": Volume (huge amounts of data), Variety (different types of media from many sources), Velocity (produced extremely fast), and Value (high usage potential).[59] Big data gives the possibility to companies, governments and others to collect, generate, and share vast amounts of data.

If used properly, big data can also be used by humanitarian and development organizations to correlate immense data sets to identify critical information for programs. For example, Data2x conducted a study that used big data generated through credit card transactions and cell-phone usage of women in Latin America to identify the priorities of individual women in the sampled areas. By analysing the information generated from 150,000 users over the space of 10 weeks, the study identified not only the amount of money spent on specific goods, but the order in which they were prioritized by women based on the order of the transactions. Using this information, researchers were able to identify the priorities of women based on their economic behavior. If conducted over the long term, this exercise has the potential to reveal how women cope following a variety of shocks, whether economic recessions, natural disasters, or conflicts.[60]

The quantity of data produced has become extremely large and diverse, making traditional data analysis techniques so time-consuming that they are not viable in some cases. Advanced computing methods provide tools to analyse huge amounts of information and identify patterns to make predictions and inform decision making. Advanced computing methods are categorized into systems that use machines, humans, or a combination of both to conduct specific tasks. This document provides key considerations to be made when human computing tasks to volunteer networks as well as when applying advanced data analytics.

Any use of big data for operations, including the decision to use or not use big data, should be coordinated with HQ by contacting DTM Analytics at DTManalytics@iom.int

Volunteer networks

Increasingly, humanitarian organizations are turning to volunteer networks to help address the enormous task of processing vast amounts of data. Volunteer networks are large groups of volunteers who apply their technical skills to support humanitarian action from anywhere in the world.[61] Volunteer networks

[59] Whipkey, Katie & Verity, Andrej. *Guidance for Incorporating Big Data into Humanitarian Operations*. Digital Humanitarian Network. September 2015

[60] *Big Data and the Well-being of Women and Girls*. Data2x. April 2017

[61] Capelo, Luis et al. *Guidance for Collaborating with Volunteers and Technical Communities*. Digital Humanitarian Netowrk. 2012

can provide additional capacity throughout the data-processing to increase the speed at which large volumes of information are processed. They may be used by IOM staff, provided the necessary precautions are taken, to aid in the processing of humanitarian data.

Examples

Crisis mapping

Crisis mapping is a participatory mapping technique to gain additional information on maps produced by satellite imagery and GIS systems. Crowdsourced crisis mapping allows the gathering of information from a large number of people, trusted informants, and publicly available sources and funnel it onto one platform.[62] Volunteers can also be crowdsourced to process and verify the information and produce tools for respondents such as crisis maps, which identify the needs and locations of affected populations.

Information management

Volunteer networks may be activated to provide real time monitoring of publicly available data sources such as social media in order to turn complex information into actionable, useful formats.[63] They are also used to validate large amounts of information, especially user-generated content.

Remote assessments

As mentioned in the section on *remote sensing*, volunteer networks can be useful in processing large amounts of data obtained through UAVs or satellite imagery. Volunteers can quickly verify images and assess damage.

Challenges

The main challenge of working with volunteer networks highlights the paradox of working with data: While making data more open increases its usefulness, it also increases the risk of problems with privacy, security and data protection.

Misuse of data

Criminal networks, or repressive governments could gain access to the uploaded information and potentially use it for their own gain, or to cause further harm to the population.

Working with volunteers

Volunteers can come from all different backgrounds, motivations and levels of experience with emergencies and humanitarian action. [64]

High-profile crises that receive the most media coverage will receive the most volunteers. For disasters that do not receive as much media attention, or become protracted problems, may make it more difficult to identify willing volunteers.

Once information is open to volunteer networks, some control is lost over how the information will be stored and handled.

[62] Chamales, George. Rob Baker, Securing Crisis Maps in Conflict Zones.

[63] http://www.standbytaskforce.org/for-humanitarian-agencies/information-management/

[64] Capelo, Luis et al. *Guidance for Collaborating with Volunteers and Technical Communities*. Digital Humanitarian Network. 2012

General recommendations Identify and train the network

The most successful crowdsourcing initiatives use groups that have undergone previous extensive training on best practices and the risks associated with data. Begin by identifying a network of trusted people ahead of time. Train the volunteers in order to understand the risks of digital information, understand needs of the organization and improve their own effectiveness. Identifying volunteer networks should be done ahead of time, previous to the response. Volunteers should understand that information should not be published beyond the group.

Identify an IOM focal point to be available at all times to the volunteer network.

Choose the tasks

Decide what tasks should be assigned to volunteer networks and what tasks shouldn't. Some things will be best handled by the organization and others not. [65] Identify the exact skills that are needed. Requests to volunteer networks should be as specific as possible, providing only the information that is necessary to complete the specified task(s).

Advanced data analytics

High capacity computers and algorithms are capable of using, combining and detecting trends and patterns in data in unpredictable and unintended manners. As many data sets become public, there is an increased risk to the privacy and security of data subjects. Data on the humanitarian impact on certain populations over time could be used in order to refine and exacerbate humanitarian harm in the future. Because this field of technology is rapidly advancing, a lot of uncertainty exists on what information could be retrieved tomorrow from datasets released today.

General recommendations

Given the uncertainty regarding information that could be retrieved from raw data[66], even if it is non-personal, using advanced analytics, it is advisable not to publish raw data in conflict areas or data on highly vulnerable populations, if not strictly necessary. Indicators on sensitive subjects should be left out to the extent possible without rendering the data useless. IOM staff will inform themselves of other data being published about a given area. As the field of **artificial intelligence** and **machine learning[67]** is rapidly developing, this guidance document will be updated regularly to take into account the latest insights. Although IOM will explore this developing space, experimenting with these technologies should always be done with adequate oversight, and IOM staff should always be the final decision maker.

[65] ibid

[66] Data that has been collected, but not yet been processed in a way that can inform decision making.

[67] These terms describe the ability of machines to construct algorithms that learn from and can make predictions on data, thus mimicking human decision-making.

Phenomena that shape risk in advanced analytics:

The mosaic effect

The mosaic effect occurs when two or more datasets are published that in themselves do not contain information, but that, when combined, reveal sensitive insights about data subjects. An oversimplified example is when a dataset about sexual assault is published without names but with container numbers, and a separate dataset is published linking container numbers with ethnicity, or even personally identifiable information.

Inference

Information can be inferred when a data set / sets contain(s) different components of information that together leave as the logical conclusion that a piece of information relates to one specific data subject or group of data subjects. The risk of inference increases as more datasets are combined.

Blind trust

Another 'phenomenon' that occurs as increasingly complex algorithms spit out 'insights', is that humans rely more heavily on these results, even though they do not fully understand the logic behind it. In cases where information is scarce, it can be tempting to rely on such insights, and it is important to be cautious of this effect.

Data Scraping

A web scraper is a program used to obtain data from websites. Data scraping software searches, downloads and extracts data from HTML formats. Scraping will typically take a portion from a page and use it elsewhere for another purpose. This could present threats to communications using HTML formats, including communications to affected populations. Websites should have measures to prevent web-scraping, such as detecting and prohibiting web bots from viewing.

Audits

As algorithms become increasingly complex, it becomes more important to subject them to regular audits, to ensure the validity of their results. Even if the algorithm itself becomes too complex for human beings to understand, its outputs can still be tested against the inputs by professionals. Especially if algorithms are used to directly inform impactful decisions, such audits are recommended where possible.

Human factor

Although insights generated by artificial intelligence are starting to inform an increasing variety of decisions in the private and public sector, IOM staff should not let these insights be the only piece of information on which to base their decisions, especially for decisions that have significant impact on peoples' lives.

Annexes

Annex 1 Existing data protection guidance from IOM and other actors

General data protection guidance

General Assembly Resolution A/RES/68/167

The General Assembly resolution called upon member states and organizations to respect and protect the right to privacy including in digital communications, to avoid and to stop any violations of individuals' privacy, to revise procedures, practices and legislation regarding surveillance, wiretaps and data collection in order to implement human rights obligation.

IOM Data Protection Manual, 2010

IOM was one of the first organizations engaged in humanitarian contexts to produce guidance on data protection. IOM's Data Protection Principles were adopted on 2009. IOM's Office of Legal Affairs at IOM Headquarters is the focal point on data protection issues and published subsequently the IOM Data Protection Manual in 2011. The application of the Manual is the responsibility of IOM departments and practitioners operating in data management in all countries where IOM operates.

The Manual defines data protection as *the* systematic application of a set of institutional, technical and physical safeguards that preserve the right to privacy throughout the various stages of data collection, storage, use and disclosure. It was based on the existing legal frameworks for personal data. It includes principles on lawful and fair collection, specified and legitimate purpose, data quality, consent, transfer to third parties, confidentiality, access and transparency, data security, retention of personal data, application of the principles, ownership of personal data, oversight, compliance and internal remedies and exceptions.

The manual also sets out a risk-benefit assessment and a sensitivity assessment to be implemented throughout the data processing cycle.

The Data Protection Manual draws on the principles of access and transparency to advocate for complaint procedures that ensure that individuals can exercise their data protection rights (e.g. right to access, rectify, and delete personal data).

Policy on the Protection of Personal Data of Persons of Concern, UNHCR

Rules and principles relating to the processing of personal data of persons of concern to UNHCR. Applies to all personal data held by the organization, but does not cover aggregated or de-identified data. Covers basic principles, rights of the data subject, data processing by UNHCR, data processing by partners, transfer of personal data to third parties, and accountability and super vision.

WFP Guide to Personal Data Protection and Privacy, WFP

Guidelines developed for WFP staff involved in the processing of beneficiary data. The document provides data protection principles as well as the application of those principles. It also provides instructions on how to conduct a privacy impact assessment.

Rules for Personal Data Protection, IRCRC

New technologies have been developed, making it possible to have increasing quantities of data. This has an enormous potential for humanitarian actors, but also has many risks to the privacy and security of data subjects. These rules are written for ICRC staff to abide by internationally recognized standards of data protection when dealing with data. The document presents ideas in the following format: basic principles, rights of data subjects, ICRC commitments, data transfers, and implementation.

The right to privacy in the digital age, OHCHR

The report verifies national legislation as well as other measures taken to ensure the right to privacy in the digital era. The report covers the protection against arbitrary or illegal interference in the private life, family, home or correspondence, who is protected and where. The report concludes that international human rights law provides a universal and unambiguous framework for the promotion and the protection of the right to privacy, while drawing attention to inadequate government compliance and adherence.

Data Protection Protocols for Crisis Mapping, ICRC

The Protocols offer a series of principles to guide data collection by protection actors. Protection actors should establish formal procedures for handling information from collection to archiving or destruction. When calling upon the general public or community to share information, actors must obtain informed consent. Actors should, to the extent possible, keep victims and communities informed about the actions taken on their behalf in addressing abuses or violations.

The Protection of Personal Data and Privacy in a Globalized World: A Universal Right Respecting Diversities, International Conference of the Data Protection Commissioners

The Conference resulted in a declaration and a call of action addressing governments and the UN. The declaration sets out 11 guiding principles for data protection reforms, namely law and fair collection and protection, accuracy, purpose-specification and limitation, proportionality, transparency, individual participation and guaranteeing the right to access (for the data subject), non-discrimination, data security, responsibility, independent suspension and legal sanction and adequate level of protection in case of a data flow across borders.

The following conference, in 2015, resulted in a resolution on privacy and international humanitarian action. The resolution included three parts. To commit the International Conference to analyse privacy and data protection requirements in the context of humanitarian action in future meetings; to endeavor to meet the demand for co-operation in the development of guidance expressed by international humanitarian actors, taking into consideration the specificities of their actions and the need to be facilitated; to mandate the Executive Committee to create a Working Group on Privacy and Humanitarian Action to lead and coordinate these activities, and to call on data protection networks to actively contribute to the working group

From Principles to Practice the Principles for Digital Development Working Group

This document details practical principles to guide the establishment of data systems. The digital principles are: design for the user, understanding the ecosystem, designing at scale, build for sustainability, be data driven, use open standards, open data, open source and open innovation, reuse and improve, address privacy and security and be collaborative.

Humanitarianism in the New Age, OCHA

This policy paper provides a comprehensive overview of the benefits of big data in humanitarian actions. It outlines four general policies that actors (governments, interagency, humanitarian actors, private sector and donors) should develop to ensure that data is transformed into useful information: ensure circulation and communication as a basic need in response; ensure freedom of sharing information, build the capacity to use information and shape robust ethical guidelines for its use, ensuring adherence to the principle of Do No Harm.

Anonymization: Managing Data Protection Risk Code of Practice, ICO

This code of practice provides clear and practical advice and explanation of legal concepts in data protection. It provides an in-depth explanation of anonymization techniques and their relevance to specific situations.

Recommendations for Standardized Implementation of Digital Privacy Controls, US Federal Chief Information Officers Council (CIO)

Rather than limiting the definition of Personal Identifying Information (PII) to data that directly identifies or can be used to contact a person (e.g. name or email address), or personal data that is very sensitive (ID card or bank account numbers), the recommendations suggest a broad, dynamic and context-specific definition that includes data elements that do not directly identify an individual but can be combined to do so. In other words, data is potentially PII if they are linked or can be linked to the specific individual. Advances in technology mean that computer programs can scan the internet for PII data elements, creating a 'mosaic' of information linking to an individual.

Guidance for Incorporating Big Data into Humanitarian Operations, Digital Humanitarian Network

This guidance is intended to assist information or data focal points in humanitarian organizations to understand the variety, categories, and possible approaches to harnessing big data for their organization's work, and includes practical checklists. Outlining potential benefits and risks, it covers internal organizational considerations, data pipeline and acquisition issues, policy considerations, contextual requirements, and the big data wheel.

Satellite Imagery

Satellite Imagery Interpretation Guide, Harvard Humanitarian Initiative

The Satellite Imagery Interpretation Guides are a series of guides produced by The Harvard Humanitarian Initiative (HHI).

SMS and mobile telephones

Towards a Code of Conduct: Guidelines for the use of SMS in natural disasters, GSMA Disaster Response, Souktel, The Qatar Foundation

Informed by existing expertise and research on best practices, this is not an all-encompassing document, but rather a guideline for rolling out SMS services during a natural disaster. It includes recommendations on the general application of SMS services, their roll-out, launch and delivery, and phase out/handover. Recommendations are addressed to responders as well as mobile network operators.

Messaging Apps

Humanitarian Futures for Messaging Apps, ICRC

Understanding the opportunities and risks for using instant messaging apps in humanitarian action. The report presents general principles for humanitarian organizations to develop strategies and standards to determine where messaging apps might be appropriate.

Social media and networks

Think Brief Hashtag Standards for Emergencies, OCHA

The OCHA Brief proposes a series of Twitter Standards in Emergencies which includes: encouraging geo-location of tweets in order to provide accurate assistance, standardizing hashtags, track information during catastrophes (three types of hashtags: a disaster named hashtag, a public reporting hashtag and an emergency response hashtag) in order to ensure the continuity of information, tracking of needs of affected population and facilitate direct assistance, promoting of the hashtag through infographics.

Report and Guidance on Privacy in Social Network Services, International Working Group on Data Protection in Telecommunications

The report focuses on how social network services have altered the definition of a person's "individual space". Before, the processing of personal data from public sources has had a privileged position in data protection and privacy legislation. However, the arrival of social network services has blurred for many users the boundary between the public and the private sphere. Therefore, this document sets out a tripartite guidance to addresses regulators, providers of social media and users.

Unmanned Aerial Vehicles (UAV)

Humanitarian UAV Network Code of Conduct, UAViators

The Code of Conduct aims to guide and support all humanitarian actors which employ UAV in order to deliver humanitarian assistance in natural disasters and situations of conflict. While the code is not binding, it does contribute to a higher standard of professionalism, safety and impact. The Code includes 14 general principles by which the humanitarian actors need to follow in order to ensure a high standard for UAV flights. These principles range from the standard humanitarian principles (neutrality, do no harm etc.) to more specific ones (i.e. be conflict sensitive, consider environmental implications and so on).

Unmanned Aerial Vehicles in Humanitarian Response, OCHA

This report will outline potential uses of UAVs in humanitarian response and emerging issues. It will also consider how humanitarians should engage with the capacities offered by UAVs used by peacekeepers or militaries in humanitarian contexts. The report will not cover the legal and ethical implications of armed UAVs or other autonomous weapons systems, although the continuing debate over their use in armed attacks will surely have an impact on the acceptance of civilian uses.

Additional guidance

IOM Migration Data Governance Policy (IN/253)

The Migration Data Governance Policy IN/253 outlines the standards that guide the Organization's Migration Data Governance. Pursuant to IOM's institutional precepts, the Migration Data Governance Policy, IN/253, ensures that IOM is principled in having a migration data governance framework for continued accountability, transparency and efficiency regarding migration data usage and sharing. This policy aligns with IOM's Migration Governance Framework (MiGoF) and IOM's Results Based Management (RBM).

The Seven Fundamental Principles, International Federation of Red Cross and Red Crescent Societies

This expression of the values and principles of the IFRC was proclaimed in Vienna in 1965. These principles are: humanity, impartiality, neutrality, independence, voluntary service, unity and universality.

Humanitarian Principles, OCHA

The humanitarian principles are the fundamental foundation for humanitarian work. The principles are: humanity, neutrality, impartiality and operational independence.

Promoting Innovation and Evidence-based Approaches to Building Resilience and Responding to Humanitarian Crises, DFID

Drawing on a review of humanitarian policy, DFID identifies four areas needing support: enabling decision makers to access and use evidence about risk, building resilience enhancing response, building capacity promoting resilience and ensuring effective humanitarian response.

Professional Standards for Protection Work, ICRC

Within a larger Its seven chapters deal with general principles, managing protection strategies, standards and guidelines, promoting complementarity, handling sensitive information and assuring professional capacity. The document touches upon new technologies with a general overview of professional standards to which actors should adhere.

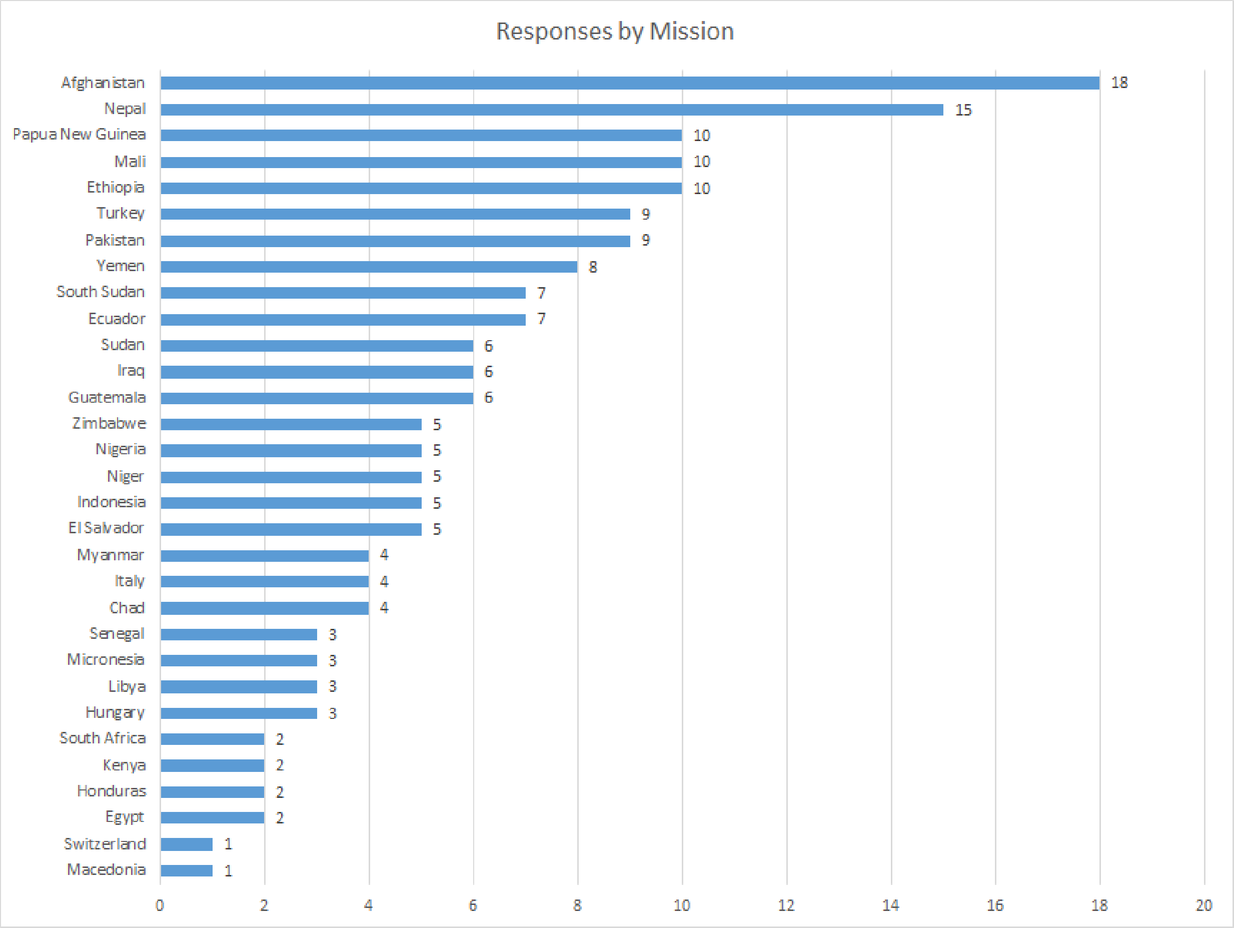
The Signal Code: A Human Rights Approach to Information in Crisis, HHI

Before developing minimum standards for using ICTs, it is important to understand that affected populations have the right to information, to data agency, to data protection, to privacy and security, and to redress their data. Only after that can organizations identify the ethical obligations and minimum standards when using ICTs in crisis.

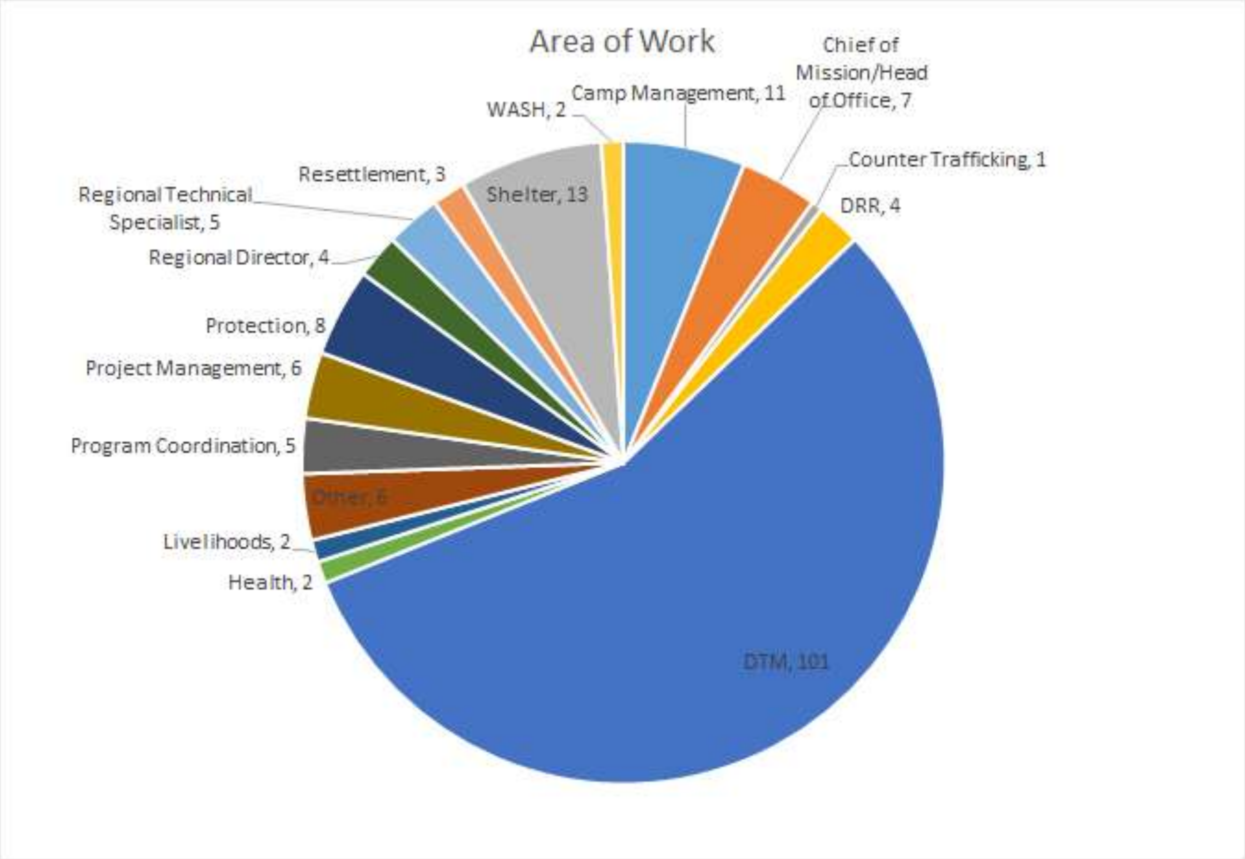
Annex 2 Staff Survey

These guidelines are the result of many requests from IOM field staff for further guidance when working with data in crises. While IOM has a Data Protection Manual and Data Protection Principles, many staff are in situations, or working with technologies for which there is no official guidance. In order to ensure any further guidance responds directly to the needs of the staff, a survey was created to be sent to IOM field offices who are conducting Displacement Tracking Matrix (DTM) activities. DTM staff work with data on a daily basis, regularly capturing, processing, and disseminating information in order to indicate the movements and needs of displaced populations. The survey allowed DTM staff to share concerns when working with data in general and when using new technologies. The results of the survey helped steer the direction and establish the scope of these guidelines.

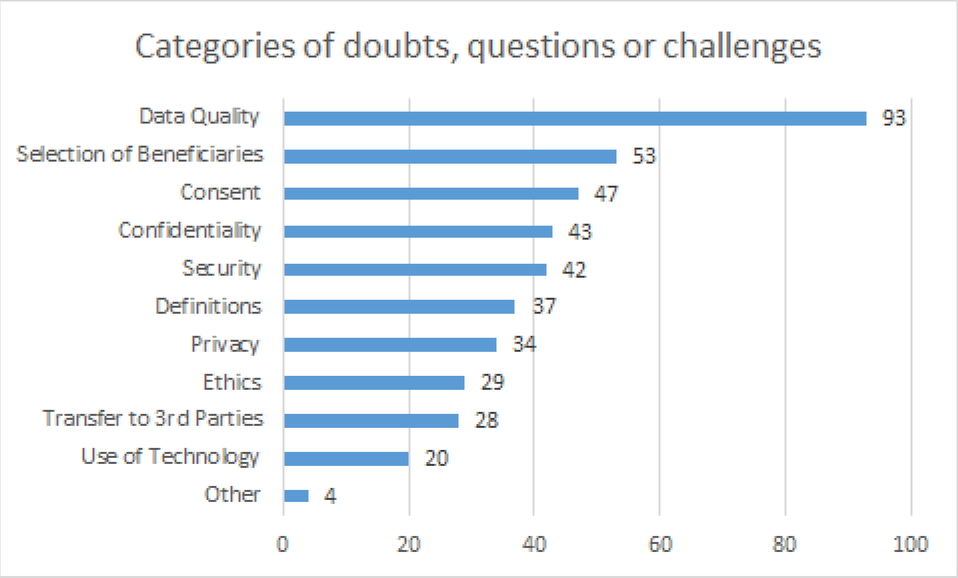
In all, 180 IOM staff from 31 IOM offices worldwide responded to the survey. Responses were received from Afghanistan, Chad, Ecuador, Egypt, El Salvador, Ethiopia, Guatemala, Honduras, Hungary, Indonesia, Iraq, Italy, Kenya, Libya, Macedonia, Mali, Micronesia, Myanmar, Nepal, Niger, Nigeria, Pakistan, Papua New Guinea, Senegal, South Africa, South Sudan, Sudan, Switzerland, Turkey, Yemen, and Zimbabwe. See figure below for a breakdown of surveys submitted by IOM office.

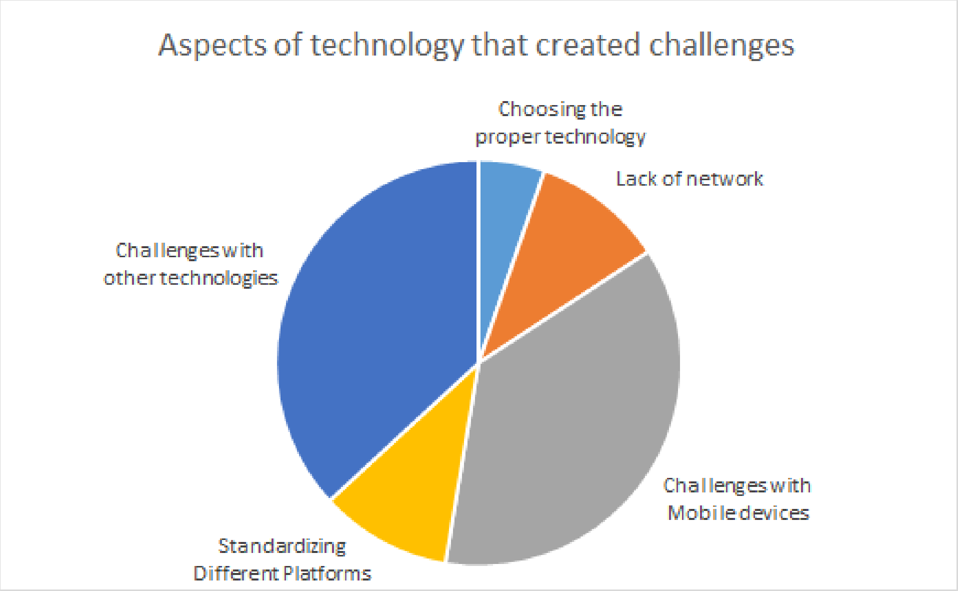


Most of the respondents were DTM staff, however there were many staff from other areas of work completed the survey as well, including those from upper-management level. This helped to capture information on IOM staff's work with data at any level including, the enumerators collecting and handling the data directly, project manager using the data for reporting, or chiefs of mission coordinating a response based on the data (see figure below). The survey quantifies the respondent's work with data: at what stage they contribute to the data processing cycle (I.e. collection, cleaning, and visualization) and identifies the purpose of the data, for whom the data is intended, and the technologies used to collect, store, or share the data.



After establishing the profile of the respondent, and their work with data, the survey then identifies general categories of doubts, challenges or questions associated with data the respondent may have. It became apparent that despite the existence of the Data Protection Manual, many felt there was a need for further guidelines for data quality, selection of beneficiaries, obtaining consent in emergencies, and maintaining the privacy and security of data subjects. In addition, many staff also expressed the need for further guidance with challenges they face when using technologies such as mobile devices and web applications used to collect and visualize the data. While this section was reserved for general categories of concern, respondents were also given a chance to give personal experiences or specific examples of any challenges or ethical issues they had faced when working with data.





These challenges reinforced the broad areas of concerns as mentioned above. In all, 71 personal experiences were given. Each respondent was asked to provide the context of the example, state the challenge, and provide information on who was affected, the technologies used, and the efforts they took to mitigate the potential problems.

Of the 71 examples, 21 affirm there are considerable issues with verifying the quality of the data. Many stated their data regarding the numbers of beneficiaries or affected persons were intentionally inflated. In most instances, the respondents were able to correct the data by triangulating the sources, verifying numbers with the government, or otherwise verifying the data with other sources. However this was often difficult in instances of emergencies as verification required heavy time investments.

Also reaffirming the broad areas of challenges or concerns, were many instances of issues staff faced while using technologies, especially the use mobile devices.[68] In some cases, the governments or members of the community did not trust the use of devices that stored information and locations out of fear that it could be used to harm them. In situations where staff were not able to gain approval to use devices from the authorities, they were left to use pen-and-paper data collection methods. Some of the following quotes from the survey highlight other issues that became recurrent themes of the challenges staff face.

"Government or other stakeholders requesting information, when their use for that information was unclear. It was feared this information could possibly be used for programming that goes against IOM's mandate."

"During the collection of the data, there were different stakeholders involved. This implies challenges in data consolidation, analysis and sharing. We were able to respond to this issue by establishing standardized procedures, protocols and standards on data collection, consolidation, analysis and sharing."

"Data presented in an IOM-bulletin was not accurate. Before the data could be adjusted, the media had already reported the inaccurate information."

While the survey did highlight challenges in data protection, it also highlighted the successes some offices had in overcoming them, for example:

"Information concerning protection issues gathered through focus group discussions was often inaccurate. We were able to resolve this by building the trust of the group through regular site visits and keeping continual contact. Once an environment of trust was established, the beneficiaries felt comfortable supplying accurate information."

The best practices and lessons learned gathered from both the challenges and successes, as described in the survey, fed into the development of these guidelines.

Copy of survey questionnaire

Working with data has become a very important part of IOM's work. While IOM has its own data protection principles and manual for guidance concerning personal data, we want to make sure that IOM operational staff have all the tools necessary to ensure adequate protection of all types of humanitarian data, especially in changing environments and new technologies. The survey below will help us identify challenges and potential ethical issues with humanitarian data faced by IOM staff. Please take time to respond honestly and thoroughly, as your answers will be used as input to the resulting guidance.

Respondent profile

Respondent Info	
Name (optional)	
Mission	
Job Title/Position	
Previous IOM missions	

[68] The use of all other technologies combined, including satellite images, UAVs, messaging apps, and social media, are included in the "other technologies category".

Educational field of study	

Under which area do you work?	
Protection	
Camp management	
Shelter	
DTM	
Counter Trafficking	
Other (specify)	

What best describes your primary role at IOM? (choose one only)			
Research		Information management	
Statistical Analysis		Coordination	
Reporting		Operations	
GIS		Other (specify)	
Database Development			

Which of the components of data management relate to your current work? (The interviewer may use their judgement to mark the best choices based on the discussion with the respondent)	
Data collection Design *(developing tools for data collection)*	
Data collection Operations *(enumerators or anyone collecting the data first hand)*	
Data cleaning *(detecting and correcting corrupt or inaccurate records)*	
Data Consolidation *(integration of data from multiple sources into a single destination)*	
Data storage *(archiving and maintaining data)*	
Data analysis *(using data to discover useful information, drawing conclusions, and supporting decision-making)*	
Data sharing and usage *(making data available to other stakeholders)*	
Data visualization *(communicating data as visual objects or graphics)*	
Reporting *(translating raw data into information)*	
N/A	
Other (please specify)	

Which type of data do you typically work with?	
Personal Data (can be used to identify an individual)	
Non-personal data	

Who collects the data you work with? (tick all that apply)	

IOM	
An implementing partner or entity working on behalf of IOM	
Third party (e.g. other organization that shares data with IOM)	
Other (please specify)	

What methods are used to gather the data you work with? (tick all that apply)			
Paper forms		Crowdsourcing or volunteer networks	
Mobile Devices		Unmanned Aerial Vehicles (UAVs or ‘drones’)	
Instant messaging apps (e.g. WhatsApp)		Satellite imagery	
Web apps (e.g. Survey Monkey)		Other (please specify)	
Social Media			

What products are made or developed using this data? (tick all that apply)	
Reports (narrative, statistical, or analytical)	
Maps (static or interactive)	
Data (raw or aggregated)	
Dashboards (static or interactive)	
Other (specify)	

With whom do you share these products? (tick all that apply)			
IOM		News and Media	
Clusters		Governments	
UN		Public	
International NGOs		Other (specify)	
Local NGOs			

How do you share personal data of IOM beneficiaries with partners?	
Agreement in place that defines how data will be shared	
Partners sign an access request form	
Nothing in writing with partners, we just share the data they ask for	
Other	
N/A	

How do you share non-personal data of IOM beneficiaries with partners?	
Agreement in place that defines how data will be shared	
Partners sign an access request form	

Nothing in writing with partners, we just share the data they ask for	
Other	
N/A	

The data is used for which of the following purposes? (tick all that apply)	
Response planning	
Aid delivery	
Coordination	
Research	
Advocacy	
Monitoring & evaluation	
Project or programme development	
Other (specify)	

In which of the following categories have you experienced doubts/challenges related to your work with data? (tick all that apply)	
Definitions (personal vs. sensitive vs. general, etc.)	
Confidentiality (protection against unintended or unauthorized access to data)	
Consent (informing respondents not just on the collection of data, but how it will be stored, processed, shared and used)	
Ethics (protecting participants from any harm associated with the data)	
Data quality (Data accurately reflects the real world)	
Selection of beneficiaries (using geographical, community, and/or household data to identify beneficiaries)	
Transfer to third parties (e.g. transfer from a remote server to a local computer)	
Privacy (how personally identifiable or sensitive data is collected, stored and used)	
Security (measures applied to prevent unauthorized access to data)	
Use of technology	
Other (specify)	

Past experiences

Do you have a specific example or past experience of a problem you faced when working with data? (i.e. doubts for potential unintended consequences related to data)	
Yes: continue	No: skip to “Current or Future Concerns”

Describe the experience and the context.			
This example was a potential problem to which of the following?			
Overall affected population		You, or your professional career	
Specific group within affected population		The government in country	
Program/project beneficiaries		Other (please specify)	
The problem in this example is related to which aspect of data management? (mark all that apply)			
Data collection Design *(developing tools for data collection)*			
Data collection Operations *(enumerators or anyone collecting the data first hand)*			
Data cleaning *(detecting and correcting corrupt or inaccurate records)*			
Data Consolidation *(integration of data from multiple sources into a single destination)*			
Data storage *(archiving and maintaining data)*			
Data analysis *(using data to discover useful information, drawing conclusions, and supporting decision-making)*			
Data sharing and usage *(making data available to other stakeholders)*			
Data visualization *(communicating data as visual objects or graphics)*			
Reporting *(translating raw data into information)*			
N/A			
Other (please specify)			
Which of the following data collection/sharing methods or technologies were used?			
Paper forms		Crowdsourcing or volunteer networks	
Mobile Devices		Unmanned Aerial Vehicles (UAVs or ‘drones’)	
Instant messaging apps		Satellite imagery	
Web apps (e.g. Survey Monkey)		Other (please specify)	
Social Media			
How serious was this particular situation, in your opinion?			
Very serious		Not important	
Serious		Don’t know	
To be considered, but not serious			
What did the mission or staff do to address the problem? (provide any measures that were in place to mitigate any potential problems or unintended consequences from the data)			

a. **Were these measures adequate?**	Yes		No	
b. **If no, what more could have been done?**				

Use the chart below to show resources that have been, or would be useful for addressing this issue. Add any that may not be listed (tick all that apply)					
	Useful	Not useful	Available	Not available	Don't know
Written standards of data protection, provided by the organization					
Guidance from my supervisor					
Guidance from my colleagues in the mission, HQ, or RO (depending on your location)					
Specific and clear data agreements "project-by-project"					
Training on data protection					
Other (please specify below)					

Do you have any other past examples or experiences that highlight potential problems that can result from data practices?	
If yes, copy questions on past experiences again	If no, proceed to next section (current or future concerns)

Current or future concerns

Are you aware of any current activities involving data that could potentially create problems? (i.e. doubts for potential unintended consequences related to data)	
Yes: Proceed	No: Skip to next section (IOM Data Protection Manual)

Describe the activity and how it could create potential problems.

This example is a potential problem to which of the following?			
Overall affected population		You, or your professional career	
Specific group within affected population		The government in country	
Program/project beneficiaries		Other (please specify)	
The problem in this example is related to which aspect of data management? (mark all that apply)			
Data collection Design *(developing tools for data collection)*			
Data collection Operations *(enumerators or anyone collecting the data first hand)*			
Data cleaning *(detecting and correcting corrupt or inaccurate records)*			
Data Consolidation *(integration of data from multiple sources into a single destination)*			
Data storage *(archiving and maintaining data)*			
Data analysis *(using data to discover useful information, drawing conclusions, and supporting decision-making)*			
Data sharing and usage *(making data available to other stakeholders)*			
Data visualization *(communicating data as visual objects or graphics)*			
Reporting *(translating raw data into information)*			
N/A			
Other (please specify)			
Which of the following data collection/sharing methods or technologies is involved?			
Paper forms		Crowdsourcing or volunteer networks	
Mobile Devices		Unmanned Aerial Vehicles (UAVs or 'drones')	
Instant messaging apps		Satellite imagery	
Web apps (e.g. Survey Monkey)		Other (please specify)	
Social Media			
How serious is this particular situation, in your opinion?			
Very serious		Not important	
Serious		Don't know	
To be considered, but not serious			
What measures are in place to mitigate any potential problems or unintended consequences from the data?			

a. Are these measures adequate?	Yes		No	
b. If no, what more should be done?				

Use the chart below to show resources that have been, or would be useful for addressing this issue. Add any that may not be listed.

	Useful	Not useful	Available	Not available	Don’t know
Written standards of data protection, provided by the organization					
Guidance from my supervisor					
Guidance from my colleagues in the mission, HQ, or RO (depending on your location)					
Specific and clear data agreements “project-by-project”					
Training on data protection					
Other (please specify below)					

Do you have any other specific examples?	
If yes, we need to copy questions on current or future concerns	If no, proceed

IOM Data Protection Manual.

Which of the following sections of the IOM Data Protection Manual do you find most relevant and useful to your work? (Select all that apply). Mark if the section provides sufficient information for your needs.

	Useful	Sufficient		Useful	Sufficient
Lawful and fair collection			Application of principles		
Specified and legitimate purpose			Ownership of personal data		
Data quality			Oversight, compliance, and internal remedies		
Consent			Exceptions		

Transfer to third parties			Don’t know		
Confidentiality			I have never read the Data Protection Manual		
Access and transparency			Other (please specify)		
Data security					
Retention of personal data					
a. What guidance, if any, do you think is missing from the IOM Data Protection Manual?					

Annex 3 DTM Data Protection Checklists[69]

The below checklists are intended to be practical tools to assist DTM coordinators to ensure compliance with IOM's Data Protection Principles. They include key factors to be taken into account at the different stages of the DTM project. Please note that the checklists are not exhaustive and you should always consult the IOM Data Protection Manual for further guidance.

The checklists should be used **prior** to commencing a DTM project.[70]

RISK-BENEFIT ASSESSMENT[71]	YES	NO
Is it clear which of the data you will be collecting under this DTM project are personal data and which are non-personal data?		
Have you made a list of all the personal data you will be collecting from the data subjects[72] under this DTM project?		
Is it clear which is the specified and legitimate objective for the collection of those personal data?		
Have you considered whether all the personal data you are planning to collect are needed in order to fulfill the purpose of the specific project? (you need to ensure that you collect the minimum personal data possible to achieve the specific purpose)		
Have you clearly identified the following roles: data controller (staff who has the overall responsibility of the personal data and who provides instructions to the data processors on how to e.g. collect, use, store, share and destroy them) and data processor (staff who process the personal data according to the instructions of the data controller)?		
Have you conducted a risk-benefit assessment prior to the collection of the personal data?[73] (At a minimum, have you listed all the risks and benefits that arise from collecting the specific type of personal data from data subjects to achieve the specific purpose?)		
Do the benefits of collecting the specific types of personal data for the specific purpose outweigh the risks?		
Have you planned to review the risks on periodic basis to identify new potential risks?		
Have you planned to check on regular basis if the benefits still outweigh risks?		

[69] These checklists are developed on the basis of the checklists included in MA/88 "IOM Data Protection Manual" and they are in line with the IN/138 "IOM Data Protection Principles". They are meant to be living documents and they may be amended from time to time. For any comments to the checklists please email sdaviot@iom.int

[70] If you have already commenced a DTM project, you can still go through the checklists to check if your project is compliant with IOM's Data Protection Principles and make changes as deemed appropriate.

[71] This checklist does not constitute the "Risk-Benefit Assessment" itself, which should be done separately. DOE and LEG are currently in the process of developing a template for conducting a risk benefit assessment. For any questions on how to conduct it for the time being please contact leg@iom.int

[72] The term « data subject » means an IOM beneficiary who can be identified directly or indirectly by reference to specific factor or factors. Such factors may include a name, an identification number, material circumstances and physical, mental, cultural, economic or social characteristics.

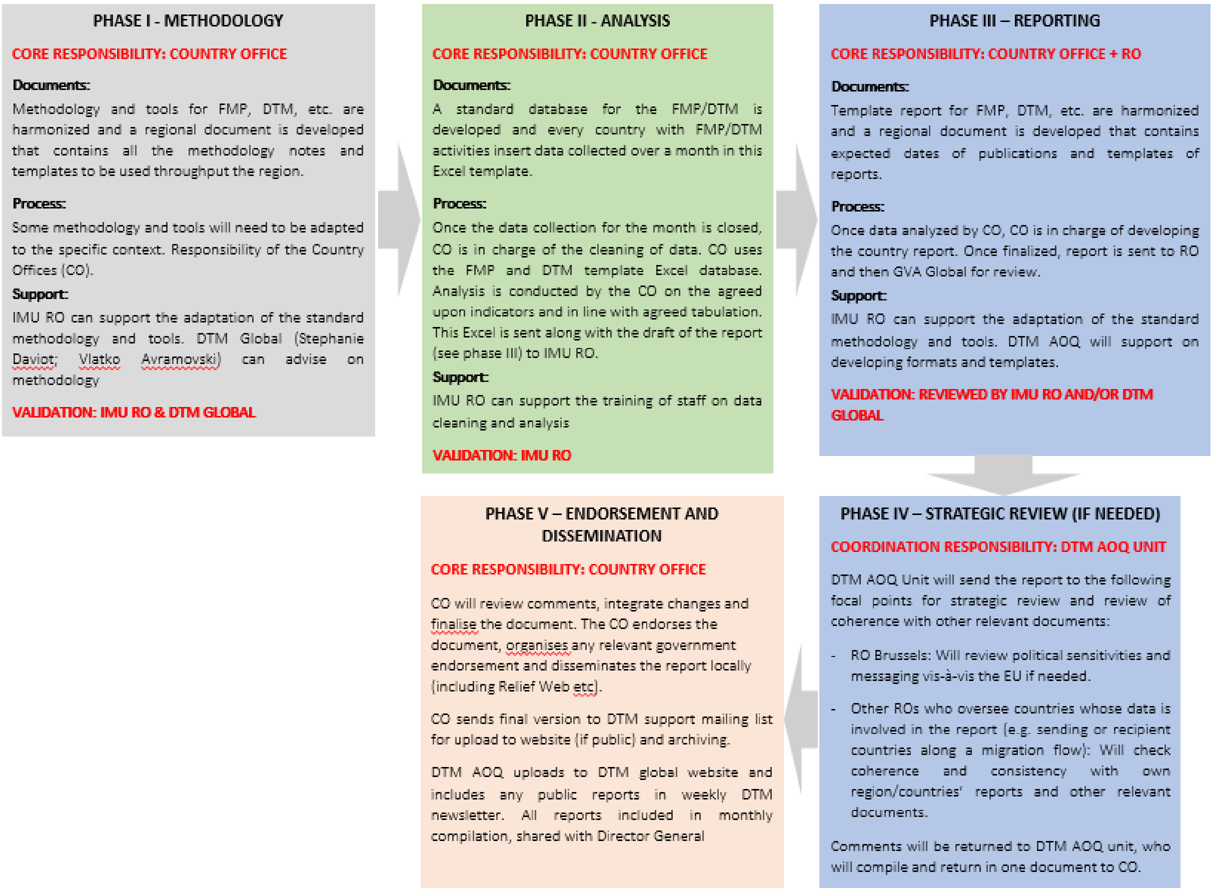
[73] See footnote 2.

SENSITIVITY-ASSESSMENT	YES	NO
Are all types personal data that will be collected properly classified according to the level of sensitivity applied to it? (i.e. low sensitivity, medium sensitivity, high sensitivity)		
Was the highly sensitive data identified? If yes, have you ensured that adequate safeguards are in place to protect such data?		
Have you planned to properly mark the personal data as being of “low sensitivity”, “medium sensitivity” or “high sensitivity” after the data will be collected?		
Have you planned to review the sensitivity of data on regular basis?		
CONSENT		
Are you able to record the consent of each data subject in writing prior to collecting their personal data?		
If written consent is not possible to obtain prior to the collection of the personal data have you evaluated the moment when such consent can be sought?		
If written consent is not possible, are you able to record the consent in another way (e.g. audio recording)?		
If recording the consent is not possible, have you contacted LEG to ensure another basis of collecting the personal data?		
Is your environment safe to seek consent of each individual IOM beneficiary?		
Is personal data collected in non-intimidating manner, with due respect of dignity of the data subject?		
Do your data subject know what the specified purpose, related purposes and additional purposes of data collection are at the time of data collection?		
Have you been providing the data subjects with an accurate and fair description of the risks and benefits at the time of data collection?		
Have you considered the data subjects’ physical and mental capacity to consent (e.g. from vulnerable data subjects)?		
Have you explained to the data subject that IOM may disclose their personal data to third parties (including donors and project partners) and have you mentioned to them for which specific reason the personal data will be shared?		
Have you explained to the data subject that he/she has the right to contact IOM to access their personal data, request to modify and/or delete them?		
Have you ensured that the data subject has all necessary IOM contact information?		
CONFIDENTIALITY		
Is your staff briefed about the confidentiality of personal data prior to the collection, use and disclosure of such data?		
Do DTM staff know that in accordance with their contracts and the IOM Staff Regulations and Rules they are obliged to respect confidentiality?		
Do DTM staff know that the obligation of confidentiality continues even after the end of their employment with IOM?		
Are you applying strict access controls to the lists containing personal data of IOM beneficiaries and maintaining an access record of personal data disclosed?		

Are you ensuring that all transmission of personal data within IOM are secure, correspondence is highlighted as “secret” and the recipients of e-mails are carefully selected?		
Are you monitoring the disposal of printed copies and other paper trails containing personal data, including the shredding of printed material containing personal data?		
DATA SECURITY		
Have you analyzed the level of security at workstations according to sensitivity levels, confidentiality, integrity, transmission and access to data?		
Have you evaluated the storage location and safety measures needed to protect paper records?		
Have you evaluated the electronic storage areas and safety measures needed to protect electronic records, including backups?		
Have you ensured the proper management of electronic and paper records to prevent unauthorized retrieval?		
Have you enquired with ICT about the latest updates in information technology , including the availability of encryption software to be used when storing personal data?		
Have you ensured a limited access to personal data of IOM beneficiaries for certain categories of IOM staff, consultants and individuals?		
Have you ensured strict access control and maintenance of personal data disclosed?		
DATA QUALITY		
Have the data subjects validated the personal data they provided to IOM?		
Have you taken reasonable steps to verify the accuracy and truthfulness of personal data at the time of data collection?		
Are the DTM staff trained on data protection?		
Are the DTM staff trained on collecting personal data?		
Has the need for truthful personal data been emphasized and have the consequences of relying on inaccurate personal data been highlighted?		
Are electronic records containing personal data stored in safe media that are protected from security risks and unauthorized access and have regular backup procedures occurred?		
Are paper records containing personal data stored in safe locations to prevent wear and tear and unauthorized access?		
Has the quality of personal data been affected by any inaccuracies?		
Have updates to the personal data been accurately recorded in the electronic and/or paper records?		
Have you encouraged the practice of cross-checking prior to collecting personal data and prior-checking before use and disclosure of personal data?		
SHARING PERSONAL DATA WITH THIRD PARTIES		
Does the donor agreement for this DTM project include a provision stating that IOM will comply with its Data Protection Principles when processing personal data?		
Have you offered/counteroffered to share aggregate non-personal data?		

Have you ensured the consent of the data subjects to share their personal data with a third entity? This is mandatory.		
Is the specified purpose for which personal data will be shared clear?		
Have you contacted LEG for advice prior to transfer of personal data? It is mandatory to sign a written agreement when sharing personal data, so you have to contact LEG.		
Have you evaluated the existence of data protection legislation, compliance with data protection laws and regulations in the country of the third party?		
Have you limited the amount of the personal data to that which is necessary to achieve the specified purpose of transfer?		
Have ensured (in coordination with the ICT Officer) that the method of transfer is safe and secure?		

Annex 4 Quality Assurance Processes for Displacement Tracking Matrix products



PHASE I - METHODOLOGY

CORE RESPONSIBILITY: COUNTRY OFFICE

Documents:

Methodology and tools for FMP, DTM, etc. are harmonized and a regional document is developed that contains all the methodology notes and templates to be used throughput the region.

Process:

Some methodology and tools will need to be adapted to the specific context. Responsibility of the Country Offices (CO).

Support:

IMU RO can support the adaptation of the standard methodology and tools. DTM Global (Stephanie Daviot; Vlatko Avramovski) can advise on methodology

VALIDATION: IMU RO & DTM GLOBAL

PHASE II - ANALYSIS

CORE RESPONSIBILITY: COUNTRY OFFICE

Documents:

A standard database for the FMP/DTM is developed and every country with FMP/DTM activities insert data collected over a month in this Excel template.

Process:

Once the data collection for the month is closed, CO is in charge of the cleaning of data. CO uses the FMP and DTM template Excel database. Analysis is conducted by the CO on the agreed upon indicators and in line with agreed tabulation. This Excel is sent along with the draft of the report (see phase III) to IMU RO.

Support:

IMU RO can support the training of staff on data cleaning and analysis

VALIDATION: IMU RO

PHASE III – REPORTING

CORE RESPONSIBILITY: COUNTRY OFFICE + RO

Documents:

Template report for FMP, DTM, etc. are harmonized and a regional document is developed that contains expected dates of publications and templates of reports.

Process:

Once data analyzed by CO, CO is in charge of developing the country report. Once finalized, report is sent to RO and then GVA Global for review.

Support:

IMU RO can support the adaptation of the standard methodology and tools. DTM AOQ will support on developing formats and templates.

VALIDATION: REVIEWED BY IMU RO AND/OR DTM GLOBAL

PHASE IV – STRATEGIC REVIEW (IF NEEDED)

COORDINATION RESPONSIBILITY: DTM AOQ UNIT

DTM AOQ Unit will send the report to the following focal points for strategic review and review of coherence with other relevant documents:

- RO Brussels: Will review political sensitivities and messaging vis-à-vis the EU if needed.
- Other ROs who oversee countries whose data is involved in the report (e.g. sending or recipient countries along a migration flow): Will check coherence and consistency with own region/countries' reports and other relevant documents.

Comments will be returned to DTM AOQ unit, who will compile and return in one document to CO.

PHASE V – ENDORSEMENT AND DISSEMINATION

CORE RESPONSIBILITY: COUNTRY OFFICE

CO will review comments, integrate changes and finalise the document. The CO endorses the document, organises any relevant government endorsement and disseminates the report locally (including Relief Web etc).

CO sends final version to DTM support mailing list for upload to website (if public) and archiving.

DTM AOQ uploads to DTM global website and includes any public reports in weekly DTM newsletter. All reports included in monthly compilation, shared with Director General