FRAMEWORK FOR DATA SHARING IN PRACTICE

PIM Protection Information Management
For Quality Protection Outcomes

FRAMEWORK FOR DATA SHARING IN PRACTICE

A. Trust Statement

There is an ethical responsibility of data and information holders to share data and information in a safe, useful manner with actors who are in a position or have a responsibility to respond to issues raised.

An environment of trust and the ways in which trust can be created, maintained, and enhanced requires working in a spirit and practice of trust, with a shared minimum approach to ensure good practice. This approach is outlined in the elements of the *Framework* below.

The trust statement will be a statement to which two parties will agree as an indication of their commitment to the *Framework* when sharing data. The statement may also extend to donors, who have the responsibility and leverage to enable data sharing and cooperation among stakeholders.

Trust Statement:

We recognize the benefits of sharing data in a responsible, safe, and purposeful manner to improve responses that promote safety, dignity, and the rights and capacities of affected populations.

We understand the risks of sharing and not sharing, and we commit to sharing and receiving data and information according to the humanitarian principles and in line with protection and information management [PIM] principles and respective organisational policies on the same.

Equipped with the Framework for Data Sharing in Practice, we will help create an enabling environment that enhances coordination and collaboration within and beyond the humanitarian community for data sharing.

If there has been a breach in trust established under the *Framework*, it is up to the stakeholders involved to understand why and the implications on the *Framework*. The *Framework* may need to be renegotiated based on the details of those terms, or it may no longer exist between the parties.

B. PIM Principles

This section summarizes the minimum shared principles that underlie and characterize the responsible handling, sharing, and use of data and information, regardless of their specific purposes, methods, or outputs (products).

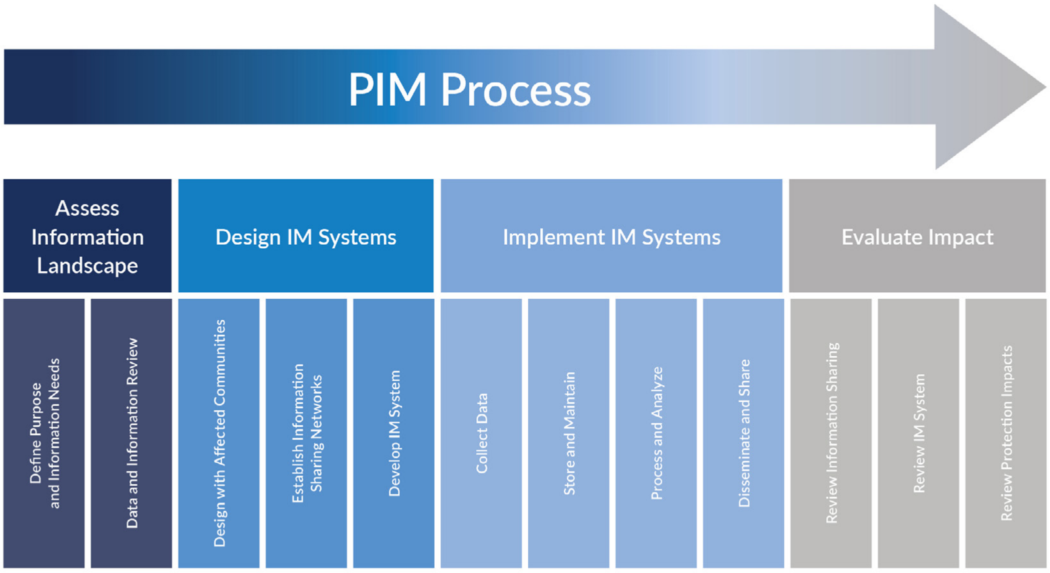
- **People-centred and inclusive:** Data and information activities must be guided by the interests, well-being, and rights of the affected population and their hosts, which must participate and be included in all relevant phases. Activities must be sensitive to age, gender, and other issues of diversity.
- **Do no harm:** Data and information activities must include a risk assessment and take steps, if necessary, to mitigate identified risks. The risk assessment must look at negative consequences that may result from data collection and subsequent actions or service delivery for as long as the data and information activity is carried out.
- **Defined purpose:** Given the sensitive and often personal nature of protection information, data and information activities must serve specific information needs and purposes. The purpose must be clearly defined and communicated; proportional to both the identified risk and costs vis-à-vis the expected response; and aimed at action

for protection outcomes, including the sharing and coordination of protection data and information.

- **Informed consent and confidentiality:** Personal information may be collected only after informed consent has been provided by the individual in question, and that individual must be aware of the purpose of the collection. Further, confidentiality must be clearly explained to the individual before the information may be collected.
- **Data responsibility, protection, and security:** Data responsibility goes beyond data privacy and data protection. It entails a set of principles, purposes,[4] and processes that seek to guide humanitarian work and leverage data to improve affected populations and their hosts' lives in a responsible manner while adhering to international standards of data protection and data security. Data and information activities must adhere to international law and standards of data protection and data security. Persons of concern have a right to have their data protected according to international data protection standards.
- **Competency and capacity:** Actors engaging in data and information activities are accountable for ensuring that data and information activities are carried out by information management and protection staff who have been equipped with data and information core competencies and have been trained appropriately.
- **Impartiality:** All steps of the data and information cycle must be undertaken in an objective, impartial, and transparent manner while identifying and minimizing bias.
- **Coordination and collaboration:** All actors implementing data and information activities must adhere to the principles noted above and promote the broadest collaboration and coordination of data and information internally between humanitarian actors and externally, with and among other stakeholders. To the extent possible, data and information activities must avoid the duplication of other data and information activities and instead build upon existing efforts and mechanisms.

C. Process for Data Management

A conducive environment for data sharing requires shared and commonly understood processes for data management, offering clear guidance on the necessary steps supporting good practice.


PIM Process
Assess Information Landscape
Design IM Systems
Implement IM Systems
Evaluate Impact
Define Purpose and Information Needs
Data and Information Review
Design with Affected Communities
Establish Information Sharing Networks
Develop IM System
Collect Data
Store and Maintain
Process and Analyze
Disseminate and Share
Review Information Sharing
Review IM System
Review Protection Impacts

Additional guidance on the shared process for data management is **available here**.

[4] Based in part on the work of OCHA's '***Building Data Responsibility into Humanitarian Action, OCHA Policy and Studies Series***, May 2016; p. 4.

Mapping and Understanding the Data Ecosystem by Context

When beginning to map available data, information, and analysis, it is important to understand the data and information management systems landscape.[5] As a first step, identify stakeholders and data information systems:

- Who has what information and data?
- What systems are in operation, where and by whom?
- What types of data, information, and analysis are these systems producing or capable of producing?

We recommend mapping the systems, stakeholders, and levels and types of data produced following the categories within the PIM Matrix[6], as this can help understand how systems and information interact, what data is being generated, and how the data or information is used.

D. Core Competencies

To process data safely and responsibly, practitioners sharing and receiving data for humanitarian response require a set of skills and knowledge, combined with the right attitude and mindset to facilitate the safe, responsible, and purposeful handling and use of data and information. Relevant minimum core competencies are **available here**.

E. Joint Benefit and Risk Assessment

The objective of the joint Benefit and Risk Assessment is to ensure that the benefits and risks of data sharing have been systematically and collaboratively assessed, and that actions have been identified to maximize benefits and minimize risks.

It is critical that the assessment ***is done jointly*** to identify all aspects of the Benefit and Risk equation. A broader understanding can inform the conditions of the data sharing, including informing the means, modalities, and frequency of the specific data sharing arrangement.

Key **questions** and key **actions**, along with supporting guidance, are outlined below. The questions are intended to guide the joint identification and assessment of the benefits and risks and the corresponding steps in a given data sharing scenario. This is followed by a list of proposed actions that can be taken to maximize the benefits and minimize the risks.

These questions and actions are indicative rather than prescriptive, and additional questions and actions could be necessary or relevant depending on context. Colleagues are encouraged to review and adapt the sample questions and actions accordingly.

E.1. STEP 1 Assess information Landscape from a Data Sharing Perspective

Q1. What do we need to know? Does the purpose of data sharing benefit the safety and dignity of affected populations? Is it critical? What are the potential negative impacts or harms of not sharing the data?

→ **WHAT TO DO: Define the purpose and potential outcomes of the data sharing exercise**

[5] The PIM Matrix can be used to assess, organise and understand an information landscape. It is available here: **http://pim.guide/guidance-and-products/product/pim-matrix-cover-page/**

[6] The PIM Matrix is available for download and use online at: **http://pim.guide/guidance-and-products/product/pim-matrix-cover-page/**

- Identify specific protection objectives (purpose) and activities to be informed by the data or information shared, and how the information will be used by the organizations it will be shared with.
- Agree on expected benefits of the sharing exercise and the alternatives and potential impacts of not sharing.
- Agree on the potential negative impacts of sharing: e.g. identify risk to people or colleagues, and legal and ethical issues related to sharing the data.
- Consider the domestic legal framework for any obligation to share information within the risk assessment process as well as potential risks that the domestic legal framework could pose for data subjects.

→ WHAT TO DO: Articulate reasons to share

- Identify and document the rationale behind the sharing by addressing the following questions:
    - What is the nature of the sharing – data collection for diverse purposes or data sharing for a new purpose? Are you sharing data through a collaborative process and purpose, or because you collected the data together but have differing purposes? Is the sharing meant to build upon existing data?
    - Are the intended uses and defined purposes for which the data or information was collected compatible? Is the intended use compatible with shared principles?

Q2. How can we maximize the benefits of data (collection) and sharing within and beyond the humanitarian sector?

→ WHAT TO DO: Work with stakeholders within the context of a trust framework

- Consult partners to identify opportunities for collaboration and data sharing. This will help to avoid duplication of data collection efforts as well as unnecessary burdens and risks for data subjects.
- Evaluate the risks and benefits of involving different stakeholders in the data or information sharing network based on the defined purpose and organizational mandates.
    - Different stakeholders will need access to different levels of details and types of information depending on defined purpose and risk.
- Consider motivations for sharing in order to further organize types of data collaborators with corresponding motivations.[7] Map stakeholders from a data sharing perspective. For the humanitarian community, motivations for sharing are primarily: Programming, response, and advocacy.
- Ensure data or information to be shared and processed is in accordance with the principles articulated in this *Framework*. Avoid using or sharing data not collected, processed, or used in accordance with fundamental data protection procedures.

[7] Based in part on the work of GovLab in this area, available at: http://datacollaboratives.org/; accessed 25 Oct. 2017.

Q3. Can the parties collecting and receiving data demonstrate the required core competencies and respect for the minimum principles and process? Do the users of the data or information demonstrate an understanding of relevant standards (PIM Principles), procedures, and policies?

→ **WHAT TO DO: Understand skills and capacities**

- Confirm that both organizations either have data sharing protocols or are a supporter of the *Framework*, as described in this document.
- Identify the relevant data protection policies and guidelines. Most organizations should have such policies and practices, which would address data protection, security, and breaches.
- Ensure partners' roles reflect technical expertise, ability, and willingness to handle and use the data competently, including managers and enumerators.

ACTIONS: Maximise *Minimum steps to maximize benefits*	**ACTIONS: Mitigate** *Minimum steps to minimize risk*
● Invite collaboration and coordinate transparently with other stakeholders (inc. affected populations) on data collection needs, locations, data collection methodologies, sharing to be able to expand geographic coverage, reduce duplication, maximize usability, and joint analysis. ● Look beyond the current parties to assess if there are other actors or benefits that could be derived from data collection/sharing (inc. affected communities, Government, and/or private sector). ● Transparence on procedures/methodology on data collection, and concerns to maximize data usability. ● Ensure that intentions of data sharing to improve outcomes/assistance for affected communities are clearly stated and understood. ● Identify the minimum data-sets or subsets of that data are required for sharing. ● Agree on feedback mechanism on how data that has been shared has been of benefit/use.	● Reduce the amount of sensitive data collected to a minimum. ● Seek consent from the originator(s) of the data where relevant, re purpose, scope, sensitivities? ● Ensure that any Data Protection Impact Assessment recommendations are adhered to. ● Review relevant legal frameworks. ● Review and implement inter-agency information sharing protocols. ● Comply with internal data protection/sharing policies. ● Capacity building of data users on PIM principles and processes and subject matter. ● Ensure informed consent from affected population if sensitive data is collected.

E.2. STEP 2 Design IM System from a Data Sharing Perspective

Q1. What is the specific data that needs to be shared? What is the level of detail and type of data to be shared (personal and/or sensitive data, trends, statistics, analysis, results)?

→ **WHAT TO DO: Assess shared data required for decision-making**

- Understand and document the intended use and purpose behind the reason to share. Reflect on and document the rationale for having the data in the first place: Was the data collected to inform any of the following:

    - Situational awareness and response
    - Public service design and delivery
    - Knowledge creation and transfer
    - Prediction and forecasting
    - Impact assessment and evaluation.
- Identify the type of data – e.g. personal, non-personal, protection data – and information to be shared as well as its associated sensitivities and risks. Thereafter, follow these steps:
    - Analyze the defined purpose alongside expected benefits and risks.
    - Reflect and document legal and ethical considerations.
    - Articulate primary data and information requirements.
- Understand the level of detailed data to be shared and distinguish between essential vs. desirable data and information through the following steps:
    - Based on defined purpose for the specific data sharing exercise, categorize possible data to be shared into at least two levels: critical data needed for decision-making and response, and non-critical data that would be nice to have for decision-making.
    - Thereafter, ensure data sharing focuses on data or information critical for decision-making.
- Document how the collection and sharing of personally identifiable data is essential to the well-being and protection of the individual(s) concerned.

Q2. What are the results of the secondary data review in regard to shared data needs?

→ **WHAT TO DO: Conduct and review the results of a secondary data review**

- Use a secondary data review to identify risks and benefits surrounding possible data or information sharing linked to certain types of data categories or data collection methods or specific points in time.
    - Document benefits and risks of data sharing in similar contexts.
    - Document existing data relevant to current purpose.
- After the secondary data review, review and revise critical data required for decision-making (i.e. update the previous list of data required).

Q3. Have you clearly defined what is sensitive and personal data in your specific context?

→ **WHAT TO DO: Understand the nature of the data being shared from a data protection perspective**

- Reflect on and document what constitutes sensitive data within the specific operational context. This is the set of information that can be purposefully or accidentally misused to harm the physical, legal, material, or psychosocial well-being and interest of the data subject (individuals, groups, communities).

  - Jointly define a classification system(s) to tailor protocols based on assessments of risk and data sensitivity by ranking the potential harm as low, moderate, or high.
  - Periodically jointly review the definition of what constitutes sensitive data or information for your context based on the specific data sharing arrangement and stakeholders involved.

NOTE: For additional details on identifying the type and sensitivity of data, see *Annex 1: Definitions and Shared Concepts*.

Q4. How have you considered present and future impacts of data sharing on the individuals and communities, as well as on data collectors' safety and security?

→ WHAT TO DO: Assess current and potential future context

- Examine the context (social, political, and security environments) from an information and data sharing perspective. Carefully consider future and possible changes in scenarios that could affect benefits and risks.
  - Be aware that data collected for a specific purpose may be used in the future for another purpose and that data may become more sensitive over time, depending on how it is used or how the security and political context evolves.
- Jointly agree and document an intended time frame for which the data or information to be shared may be used.
  - Plan to periodically review whether information or data needs changed from the original purpose for which the data was collected. If so, a new purpose may need to be jointly defined and applied.

Q5. If personal data is to be shared, was informed consent obtained for the intended purpose?

→ WHAT TO DO: Review data collection processes on informed consent

- Information cannot be used for purposes that are incompatible with the purpose for which it was originally collected, and in the case of personal data, for which consent has been received.
  - Identify limitations with previously collected informed consent with regard to defined purpose.
  - Data should not be used or shared for other purposes without additional consent. A further assessment of the risks and benefits would be required if shared data is to be used for a new purpose(s).

Q6. What are the needs for the data sharing arrangement, and are they appropriate to the context?

→ WHAT TO DO: **Establish data sharing arrangements and processes**

- Document and maintain an analysis of the type of data to be shared within the network, articulating what form, frequency, through which channels, and at what level of aggregation. Track who will have access to the shared information and for which purpose.
- What are the actual practices within the information sharing network (e.g. self-censoring, informal decision-making on how to share data, etc.).
- Include 'red lines' or thresholds that may require remedial or mitigation measures or the cessation of processes or activities.
- Make an informed collective decision on how to proceed with the given data or information sharing arrangement.

ACTIONS: Maximise

Minimum steps to maximize benefits

- Ensure that the methodology is fit for purpose, proportional, and transparent.
- Ensure transparency on the history of the data (traceability) by attaching metadata in a standardized way to all datasets.
- Identify all actors who can utilize the data to benefit the affected pop (including how and in which ways they will do so).
- Establish protocols and scope (including technical methods / approaches to ensuring meaningful & safe access to the data).
- Build in feedback loops for the affected population and members of the information sharing network.
- Use HXL and other relevant standards for interoperability and take all reasonable action to prevent redundancy (including in sharing).

ACTIONS: Mitigate

Minimum steps to minimize risk

- Establish mechanisms for rectification and redress for affected populations regarding their data.
- Establish clear protocols and mechanisms for scenarios of data breach (including informing relevant parties and documenting critical incidents).
- Establish and adhere to a data retention plan.
- Establish and socialize SOPs for information sharing network articulating organizational and individual accountability, roles & responsibilities throughout data lifecycle.
- Reflect the sensitivity of the data in the design of the information sharing mechanism, recognizing policy and legal frameworks.
- Train practitioners (data users and providers) on responsible data practice.

E.3. STEP 3 Implement IM Systems

Q1. Have new benefits or risks emerged in the implementation stage? Have the prevention and mitigation actions identified in Step 2 ('Data and Information Review') been successfully implemented?

→ WHAT TO DO: **Actively monitor for benefits and risks during and after data sharing takes place**

- Review data sharing plan with data sharing partners and document the following:
    - Does the implementation follow the agreed design and plan? If not, when in the process did the variance occur and why?

  - Are revisions needed to the data sharing arrangement to improve sharing or remedy issues?
- Review and document benefit or harm resulting from the sharing.
  - Has the purpose and/or environment changed (time, context, capacity) in way that may impact the data sharing arrangement?
  - Have all parties clearly communicated – including to the affected communities – benefits and risks as a result of data sharing?
- Be attentive to signs that the data sharing activity may create risks or that mitigation measures may be failing.

ACTIONS: Maximise

Minimum steps to maximize benefits

- Understand and invest in necessary technical, human, and institutional capacity, competency, and capability required to implement the data sharing approach.
- Use a phased approach to implementing data sharing modalities, especially with 'new' systems and/or approaches.
- Where possible, periodically review and adjust implementation approach and IM system based on experience.
- Ensure that necessary framing and context is provided whenever data and analysis are shared to support accurate and beneficial interpretation + use.
- Document how data is used; share within information sharing network.

ACTIONS: Mitigate

Minimum steps to minimize risk

- Document risks and harms realized; support critical incident reporting / monitoring and tracking.
- Take corrective action whenever harms are realized; document this action toward a systematic approach.
- Offer a variety of training and capacity development activities related to standards, procedures, and policies for data sharing.
- Conduct regular audit/review of users capacity and system use.

E.4. STEP 4 Evaluate Impact (of Sharing)

Q1. What were the impacts of the data sharing?

- Review and document information sharing:
  - Was it more or less than sufficient for the purpose?
  - Were the intended benefits achieved?

Q2. Have you been able to evaluate the data sharing arrangement?

- Review and document the actual data sharing process and procedures
  - Was the identified information shared as planned?
  - Were risks mitigated effectively?
  - Were there any benefits and/or risks that you did not anticipate? How were these addressed?
- Document and agree on the relevant resources (funds, people, etc.) to support sharing,

and identify and jointly advocate for more or less resources (system, people, money) in the future.

Q3. Was information shared with the affected populations as planned? What was the feedback, and how was it considered regarding their use of and access to the information?

→ **WHAT TO DO: Ensure responders have the information they need (including affected populations)**

- Work with the community to vet incoming requests for data and information. This process will need to ensure that these requests are unbiased. They also will need to focus on a commonly agreed defined purpose that is proportional and will deliver protection results on behalf of persons of concern while also reflecting agreed principles.

ACTIONS: Maximise

Minimum steps to maximize benefits

- Feedback to data providers on the usefulness and impact of using the data, including a dialogue re quality and relevance of data.
- Seek feedback from the affected population on the availability and usefulness of information, and the means of sharing that information.
- Acknowledge (human/financial) resources and sources of original data, including recommendations for fund raising, institutional capacity/competency (appropriate attribution).
- Implement the findings and recommendations of the evaluations (to adjust the sharing framework).

ACTIONS: Mitigate

Minimum steps to minimize risk

- Be transparent around bottlenecks and issues to be addressed to rectify data sharing issues - inc roles and responsibilities.
- Acknowledge unforeseen risks and mitigation actions and work towards improving/resolving them in a coordinated and collaborative manner in an environment of trust.
- Work towards expanding the inclusion of stakeholders in the data sharing framework - including by showing positive case studies.

ANNEX Shared definitions and concepts

The objective of this section is to define a minimum set of terms to facilitate a normative discussion on data and information sharing and the *Framework*.

What minimum responsibilities do data providers have for secondary uses of data?

For details, please see Section E, Joint Benefit-and-Risk Assessment.

What is risk in the context of data and information sharing?

Risk will vary by context, time, and people involved. It may include immediate or delayed threats to life, security, identity, or freedom, individually or collectively.

The proposal to address complexities surrounding risk will focus on how to begin to build a profile of risk that is specific to the situation at hand, based on defined purpose and intent behind the data to be shared. The goal is for risk to be monitored/identified and mitigated in an ongoing manner.

For additional details and considerations, please see Section E, Steps 1 and 2, which deal with defined purpose as well as context and temporality.

What is personally identifiable vs. non-personally identifiable data?

The current definitions of personally and non-personally identifiable data continue to be challenged by modern technology. Because we are dealing with human data, all of it may carry the risk of being personally identifiable.

We recommend focusing instead on how to prevent harmful use and how to assess risk through the operationalization of a shared risk-and-benefit assessment. Doing so can clarify the actions needed to assess or prevent risk, based on a series of steps. The shared analysis of the risks and benefits for a particular data sharing process would then be the component of this process that is shared.

What is informed consent?

'Informed' implies that the data subjects should receive explanations, in simple language, on the identity of the data collector or other actor and the purpose, scope, method, intended use, and potential risks of the data provision as well as the meaning of confidentiality. 'Consent' signifies the data subject's voluntary approval for the information to be used or shared as explained.

Thus, informed consent is voluntarily and freely given based upon a clear understanding of the facts, implications and future consequences of an action. According to the circumstances, it can be verbal, written, or otherwise provided according to *Best Interest Determination*[8] procedures on behalf of a minor.

Consent is regarded as freely given only when the data subject has a genuine choice and can refuse or withdraw consent without detriment.

For additional details, please see Section E, Step 2, on 'informed consent'.

[8] | *Convention on the Rights of the Child*, Committee on the Rights of the Children, GC No. 14 (2013), on the right of the child to have his or her best interests taken as a primary consideration (art. 3, para. 1), available online at: **www2.ohchr.org/English/bodies/crc/docs/GC/CRC_C_GC_14_ENG.pdf**, accessed 9 April 2018.

What is personal data vs. non-personal data?

Personal data, also known as personally identifiable information (PII), is data relating to an identified individual or a person that can be identified from that data, from other information, or by means reasonably likely to be used related to that data. This could include, for instance, an identifier such as a name, an identification number, location data, audio-visual material, or an online identifier. Personal data also includes: country of asylum, individual registration number, occupation, status, religion, and ethnicity. And it includes biometric data such as a photograph, fingerprint, facial or iris image, as well as any expression of opinion about the individual, such as an assessment of their legal status and/or specific needs.[9]

What is sensitive data vs. non-sensitive data?

Sensitive protection data and information is data or information whose disclosure or unauthorized access is likely to cause:

- Harm (such as sanctions, discrimination, repression, or stigma) to any person, including the source of the information or other identifiable persons or groups; or
- A negative impact on an organization's capacity to carry out its activities, including due to reputational damage.

Sensitivity of data is defined in relation to the particular context and level of aggregation, and may change over time. Therefore, the same data may not have the same level of sensitivity in different contexts.

Protection data and information that does not contain personal data may nevertheless be sensitive. It may relate to communities and other groups,[10] anonymous individuals, or specific events or issues. In armed conflicts and other situations of violence, various aspects relating to the humanitarian, human rights, political, or security situation may exacerbate the risks to people.

Likewise, aggregated or pseudonymized[11] data may still be sensitive. Individuals or groups may still be identifiable, especially depending on the location and sample size, and thus may be exposed to harm if data about them is disclosed.

It is therefore, not possible to propose a definitive list of what types of data or information constitute sensitive information. However, some key types of information may belong to this category, including information about the nature of violations affecting specific individuals or groups, details about victims and witnesses, the affiliation of perpetrators, operational details related to military operations or security, etc.

Privacy, security and integrity of individuals or groups may be put at risk even if no personal data is collected and processed. Recognizing this, protection actors as a matter of best practice apply the standards derived from the principles of data protection to sensitive data and information used for protection purposes, to the extent that it is necessary given the particular sensitivity of the data.

The below diagram from the Professional Standards for Protection Work (2018, page 112)[12] illustrates the relationships between types of data and information:

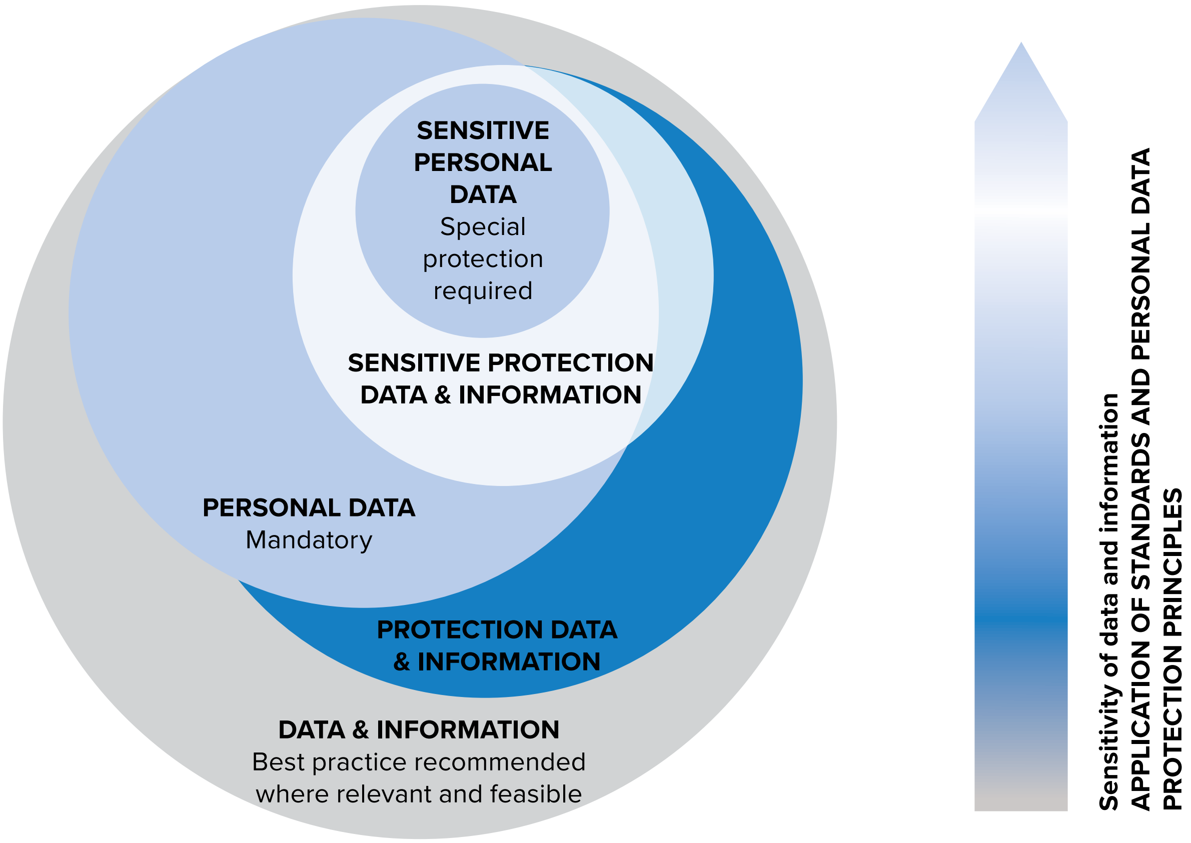
[9] | *Professional Standards for Protection Work*, *Chapter 6*, 2018.

[10] | Such data or information may be referred to as 'community identifiable information' or 'demographically identifiable data (CII / DII).

[11] | *Professional Standards for Protection Work, Chapter 6*, 2018: 'The "Pseudonymization" of data means replacing any identifying characteristics of data with a pseudonym, or, a value which does not allow the data subject to be directly identified. For example, "Jane Doe" could be pseudonymized to "POC 15364". Pseudonymization should be distinguished from anonymization, as it only provides a limited protection for the identity of data subjects in many cases as it still allows identification using indirect means. Where a pseudonym is used, it is often possible to identify the data subject by analyzing the underlying or related data.'

[12] | *Professional Standards for Protection Work, Chapter 6*, 2018.

Diagram: Relationships between types of data and information


SENSITIVE PERSONAL DATA
Special protection required
SENSITIVE PROTECTION DATA & INFORMATION
PERSONAL DATA
Mandatory
PROTECTION DATA & INFORMATION
DATA & INFORMATION
Best practice recommended where relevant and feasible
Sensitivity of data and information
APPLICATION OF STANDARDS AND PERSONAL DATA PROTECTION PRINCIPLES