CHECKLIST OF KEY CONSIDERATIONS TO KEEP IN MIND IN ORDER TO SELECT A NEW DIGITAL DATA COLLECTION TOOL IN A RESPONSIBLE WAY

The health crisis has prompted many organizations to **adapt their data collection practices**. These changes included **using new tools**, or being forced to **scale up tools that they had only previously used on a smaller scale**. For instance, some organizations began collecting sensitive data such as Sexual and Gender Based Violence (SGBV) alerts using CATI (Computer Assisted Telephone Interviewing) software, when they had only previously done so through paper-based data collection. Others generalized the use of smartphones to all their data collection activities, which until then had been limited to monitoring infrastructures.

The introduction of new tools or the scaling up of these tools necessarily raises **data protection issues as soon as the data collected is personal and/or sensitive**. However, it often turns out that local and international aid organizations do not have the resources (time, human resources, etc.) or even the skills to **analyze tools from a data protection perspective**.

This checklist therefore aims to provide some elements of popularization and awareness raising to help in the decision-making process of choosing a new data collection tool through a data protection lens.

The issue of data protection remains a complex one and this checklist only intends to make these queries a little bit more "accessible". **Under no circumstances can its use replace in full neither a detailed legal analysis nor a DPIA** (Data Protection Impact Assessment) that organizations have in many cases the legal and ethical obligation to conduct before introducing a new data processing tool and/or before initiating a new data collection process.

This checklist also does not include the following wide-ranging and essential questions related to the use of these solutions: What is the legal basis for data processing (e.g. informed consent)? How can the sensitivity level of the collected data be assessed beforehand? How to ensure a minimal and proportional collection of data according to the existing information needs? What internal procedures for data collection (design or adaptation) exist? How can teams be trained about responsible data management principles? How to configure and secure mobile terminals (encryption, PIN code, etc.)? And so on and so forth.

- For more information on these issues, please consult the various existing and specialized resources in the sector such as ICRC's "Handbook on data protection in humanitarian action", the Guidance Notes of the Centre for Humanitarian Data, the feedback from OXFAM on Responsible Data, etc.

Finally, **this checklist focuses on generic tools intended**, for example, **for household surveys**; it does not deal with the extremely complex issues of "contact tracing" and other dedicated medical applications.

CHECKLIST

To keep the checklist as straightforward as possible **the generic word "tool" will be used below to refer to a digital data collection solution**. However, this term covers very different situations, ranging from an open source IT solution that can be installed on an organization'sits own server (such as ODK Central), to a proprietary system offered as a SaaS (Software as a Service) by an independent service provider (such as SurveyCTO offered by Dobility), to solutions made available free of charge by actors of the sector (such as KoBoToolbox by OCHA). However, each of these situations covers different IT and legal realities that each requires a specific analysis that will only be briefly mentioned below to keep the checklist accessible to a wide range of aid staff.

1 Does the tool offer any initial guarantees/commitments on its approach and on the measures it implements to ensure the protection, confidentiality and securing of the data it collects and stores?

- Is it explicitly mentioned that the tool can be used for personal and/or sensitive data?
- Does it come with a clear and understandable privacy policy (or general conditions) available [and not written in an intricate legal vocabulary]? And/or does one have access to a regularly updated page summarizing the existing security measures that are being implemented?
- Is the ownership of the data collected and stored guaranteed in the contractual documents? In other words, is the potential service provider prohibited from reusing the data for any other purposes (targeted advertising, behavioral analysis, etc.) or are there any uncertainties regarding this particular aspect?
- Is the tool known to have already suffered from data breaches or to have had practices that are not compatible with the sector? Is it banned by other aid organizations?

2 Does the tool offer the required legal guarantees (particularly in the case of SaaS or tools available for free)?

Legal guarantees must - generally speaking - be compatible with the legislation of **both the organizations' headquarters** (GDPR for an organization based in Europe for example) **and each of the countries in which the data is collected** (Data Protection Act for data collected in Kenya for example); keeping in mind that some legislations may not be compatible with others.

- Do the terms and conditions specify "basic" elements to ensure compliance with the applicable legislation such as: (i) the existence of a DPO contact or a similar position, (ii) the guarantee of a maximum period to inform the relevant parties in case of a data breach, etc.?

- Do the terms and conditions specify "basic" elements to ensure compliance with the applicable legislation such as: (i) the existence of a DPO contact or a similar position, (ii) the guarantee of a maximum period to inform the relevant parties in case of a data breach, etc.?
- Does the location of the servers comply with legal requirements? For example, a server must be based in Europe or in a country considered as adequate under the terms of the GDPR.
- Do the contractual provisions comply with the law? Very often, specific DPAs (Data Processing Agreements) must be signed for the organization to be in compliance.
- Or can the tool be hosted on the organization's own servers?

3 Does the tool allow the organization to easily handle users who need to contribute to, or simply access the data?

- Is it possible to set up individual accounts [rather than generic accounts that require sharing a password between several users]?
- Can an administrator easily get an overview of all users and their respective rights?
- Can an administrator easily manage these users and their rights (delete their account or disable their access, etc.)?
- For ease of use, can the tool use the same user control and authentication system as the one already used within the organization (SSO Single sign-on with AD/LDAP)?
- Does the tool allow the configuration of sufficiently granular data accesses? Are there different rights (read, modify, delete, etc.) per user? Are these rights configurable for each data collection or type of data collection (by type of data collection activities and/or geographical areas and/or teams and/or fields)?

4 Does the tool allow to easily identify the data collection activities presenting risks?

- Does the tool allow you to manually categorize a collection activity as containing non-sensitive personal data, sensitive data, identifiable demographic information (IDI), etc.?
- Does the tool include a functionality which can automatically identify data as being personal (e.g. through the existence of a name, address field, etc.)?
- Does the tool allow for different accesses to different types of fields for the same data collection (pseudonymized data fields accessible to some users and personal data fields accessible only to others)?

5 Does the design of the tool contribute to the implementation of good data practices?

- Does the tool allow an adapted collection of consent according to the context (e.g. audio recording, via a signature, automatic printing of a receipt and subsequent contact modalities, etc.)?

- Does the tool allow for easy manual deletion of data both at the collection activity level and at the individual record level (mass deletion possible)? Does the tool allow automatic deletion of data stored locally on the terminals?
- Does the tool make it easy to track data storage periods: for example, is it possible to assign to each collection activity an expiration date with automatic reminder for data deletion? Is there an archiving function (with limited access) available?
- Does the tool allow easy pseudonymization of data?
- Does the tool warn the user when performing "hazardous" operations such as exporting data or sharing data publicly?
- Does the tool make it easy to address specific requests from a data subject about his or her data (access, deletion, limitation of the processing, etc.) by making it easier to search through certain fields (name, address, etc.)?
- Is the encryption of certain activities or data fields easy to set up, so that it does not discourage the user? Are data exports also automatically encrypted?
- Can the tool automatically anonymize certain fields such as blurring people's faces in case of photo collection, de-identification of GPS coordinates, etc.?

6 Does the design of the tool facilitate the work of the tool administrator?

- Does the administrator have access to a centralized event log (or audit trail) that makes it possible to identify the actions of each user (viewing, downloading, modification, etc.) and to identify problems in the event of a data breach?
- Does the administrator get an alert in the event of suspicious user behavior (downloading large amounts of data, unsuccessful access attempts, etc.)?
- Is the administrator helped in his user management (automatic deactivation for example of inactive users after a certain period of time, possibility of transferring rights easily in case of staff being on vacation or HR turnover)?
- Is the administrator helped in his management of collection activities (identification of activities that need to be deleted because they have expired, list of activities that are publicly accessible, etc.)?

7 Is the tool secure enough to host personal data?

- Is end-to-end encryption possible both at the server level (at rest), during data transfer (in transit) or at the terminal level (mobile or browser)?
- Is the type of encryption used known and satisfactory?
- Is the encryption key held strictly by the organization and not accessible by the service provider?
- Are the provider's servers well certified (ISO, NIST, HDS, HIPAA, etc.) with an adapted monitoring system?

- Is an external evaluation of the security measures deployed by the tool and/or the service provider available (tool security audit report, open source tool whose code has been reviewed by a community of developers, etc.)?
- Are updates of the tool carried out on a regular basis?
- Are the authentication measures compatible with those of the organization (strong password, automatic logging off after a period of time, two-factor authentication for certain users, possible access through a VPN)?
- Are the servers storing the data automatically replicated/backed up at an appropriate frequency? Does the service provider have a business continuity plan? Is the level of service (SLA - Service Level Agreement) such as server uptime or time to patch a vulnerability adequate?

RESOURCES AVAILABLE ONLINE

1 **Detailing generic data protection principles for mobile data collection:**

- **Data Security with Mobile Forms**

2 **Containing some elements of comparison between different tools about data protection:**

- **How to Choose a Mobile Data Collection Platform**
- **Benchmarking of Mobile Data Collection Solutions**
- **What electronic tools are appropriate to meet the needs of outpatient programs of Médecins Sans Frontières?**

Thank you for reading!

The icons used in this document were created by DinosoftLabs, Eucalyp, Freepik, monkik and Smashicons, available on www.flaticon.com

This project was co-funded by the French Development Agency (AFD) and the H2H Network's H2H Fund, the latter supported by UK aid from the UK government. Nevertheless, the ideas and opinions presented in this document do not necessarily represent those of the H2H Network, UK aid or AFD.