ODK AND DATA PROTECTION

Summary

Data Protection is important for our work. This handout includes overall questions on how these will affect data and information workflows. This is a draft on key recommendations and some basic research. Please edit.

Recommendations

1. Review and update Standard Operating Procedures: By writing out the Standard Operating Procedures for ODK, Kobo etc, we will be closer to meeting the new guidelines. A policy is not enough anyway. We also need a proper workflow analysis (gaps, risks etc) to back-up the SIMs needs.
    a. Every data set should have a 'version control and handoff procedure' (light weight). This will offset 'reuse of dataset' inquiry.
2. Training and Data Protection Guidance: It might be helpful to have a shared training or guidance document for all data and information workflows on data protection, even if we all work in different countries. This will show preparedness. E.g. the UK office gets only aggregated data via excel (email) spreadsheet from x deployment.
    a. Also include guidance on de-identification of personal data, using pseudonymisation (masking) or anonymisation (aggregation, conversion, etc) of dataset. Other examples: Still images of an individual or community, video footage of an individual or community, DNA samples of an individual or community, and social security numbers.

RISKS	PRIORITY RATING /MITIGATION	BACKGROUND LINK(S)	NOTES
Identify clear data governance: who is responsible for collecting, storing, processing and releasing of personal data in the organisation?			
Tech setup -security, hosting, storage			
Lack of mapped technical workflow(s) to meet RESPONSIBLE DATA USE guidelines			
Data collection processes - consent, data minimization/mvp data set			
Collected data contains personally identifiable information (PII)			

Collected data contains Demographically Identifiable Information (DII)			
Missing business / legal analysis on risks and preparation for RESPONSIBLE DATA USE			
Data controller does not manage this process			
Analysis on proprietary tool includes full data set			
Processing guidelines			
Processing by NS			
Processing by IFRC			
Processing by outside actor and out outside tool (academics/businesses)			
Transferring/Sharing - internal (EU)			
Transferring/Sharing- internal (Non-EU)			
Transferring/Sharing - external			
Inability to reach most vulnerable/do data and information workflows by adhering to Responsible Data Use/Data Protection regulations			
Disaggregation/Re-aggregation			
Archiving			

Research Background

ODK Docs
http://docs.opendatakit.org/security-privacy/

Risks

1. None of the downloadable ODK software transmits or communicates any information back to us and the software we have written does not have any mechanisms that might allow us

to access or control your devices or systems.
2. There is always the possibility that hackers can discover and exploit deficiencies or bugs in our software or in 3rd-party libraries to access or control your devices or systems.
3. When setting up your own webserver to run ODK Aggregate, if you do not configure the server and ODK Aggregate to use an SSL certificate, a determined observer can see all data communicated to and from that server.
4. With all 3rd party hosting services, you should expect your data to be viewable by the support staff of the hosting service. Different services go to differing lengths to restrict access to, encrypt, and/or secure the data and communications within their data centers.
5. We gather anonymous aggregate user behavior through Google Analytics. We use secure HTTPS communication to transfer this data off the device and the data are available to ODK's maintainers. Users may disable analytics in the settings of ODK Collect.

Handbook on Data Protection (ICRC, chapter 11)
https://www.icrc.org/en/publication/handbook-data-protection-humanitarian-action

Key risks
- need for clear guidance on processing by humanitarian organizations of information extracted from the messaging app
- lack of awareness about the types of data they process
- metadata could be accessed and analyzed by third parties and used by them in ways detrimental to the vulnerable populations.

Towards a Secure Framework for mhealth - case study in mobile data collections
http://bora.uib.no/handle/1956/10652;jsessionid=54B69ECBBED3CE2D9421018B2FD64481.bora-uib_worker
"For this work, we collaborated with the open-source MDCS, openXdata and Open Data Kit (ODK)."

P. 31 case studies considering security and privacy in data collection, transfers and archives. Very clear explanations of risks and security workflow gaps in international development

General Mobile data
http://blog.safedk.com/gdpr/gdpr-meaning-mobile-apps/
To ensure that data processors can accurately create a complete history of change while guaranteeing confidentiality, the following measures must be implemented in mobile app design, installs and usage:
1. **Determine whether the app really needs all of the data**
2. **Inform the user and obtain consent**
3. **Respond to user requests**
4. **Encrypt user data**
5. **Ensure users are updated about security incidents**
6. **Know your technology and potential weak links**

- Identify and study all relevant processors to understand what data is stored and processed, how well each processor protects personal data, and how they are working towards a responsible data policy.
- Study your own data locations and practices and ensure the data is segregated and protected.
- Ensure you have a strong internal security policy and enforce it.
- Determine whether you need to hire a DPO.
- Inquire and make sure that the SDKs you work with don't gather and save data in their own databases.
- Map out the path the data takes during the processing lifecycle to ensure adequate security is implemented at each stage.
- Make sure that the SDK has adequate security measures to ensure the safety of your

users' data. Include strict confidentiality, data privacy and data residency clauses in any contract drawn up with an SDK.
- Take advantage of automated tools that can help you stay in control and consistently monitor 3rd party providers' impact on your app, as well as provide alerts and help you take measures to handle problematic areas.

Privacy Implementation Assessment template
http://docplayer.net/63464342-Towards-a-privacy-impact-assessment-template-for-mobile-health-data-collection-systems.html

Privacy IA template is typically structured in four parts:
1.Description of the application (i.e., Mobile Data Collection Systems - MDCS), in terms of its objective, requirements, users, stakeholders, application arc
hitecture, and data flows.
2. Identification of privacy threats with respect to a list of privacy targets embedded in the legal framework EU GDPR.
3. Proposal of technical and organizational controls for mitigating the identified privacy threats, i.e., counter measures
4. Documentation of the PIA regarding the MDCS being analyzed

Threat identification and selection of countermeasures
By looking into the MDCS context, a list of privacy threats can be postulated and associated to the aforementioned privacy targets. This threat identification analysis can be carried out by a group of experts, using brainstorm sessions and iterative reviews.
Two preliminary examples of such threats are

Threat 1: User profiling (data minimization threat) The use of MDCS makes it fairly easy to link subject's data (i.e., patients or families), activities, kinship, demographics, and etc. User profiling is inherent in the health surveillance process. The further exploitation of data
relationships, creation of more complete profiles, might result in the use of personal data beyond the original purpose.

Threat 2
Vague purposes (purpose biding threat): Vaguely defined purposes allow MDCS to be used for purposes not previously defined during the design stage. For example, do not follow the premises of meaningful use of medical data, and uses it for secondary purposes