TIP SHEET 7

DATA RETENTION, ARCHIVING AND DISPOSAL

elan
The Electronic Cash Transfer Learning Action Network

WHAT ARE DATA RETENTION, ARCHIVING AND DISPOSAL?

Data retention refers to the length of time that data is kept by the organization that gathered it. **Data archiving** describes the intentional preservation of data in a format that makes it easy for collaborators to refer back to. And **data disposal** is the process of deleting data in a safe and responsible manner. All are critical to safely and securely managing program data.

WHAT HUMANITARIANS NEED TO KNOW:

Plan from the outset

Although retaining, archiving and disposing of data all occur at the end of the data and program lifecycles, plan for them from the beginning. Leaving these decisions to the end of a program increases the risk that they will become rushed afterthoughts. Without deliberate planning, you may run out of resources or time to responsibly manage this stage of the data lifecycle. This lack of planning can put program participants at risk or cause you to store data longer than necessary.

Understand the needs of all stakeholders

Even within the same program, different teams can have different data retention needs. For example, finance and compliance staff might need to access program data for a donor audit years after a program ends. In contrast, program staff may no longer need data once the program is complete. To meet the data needs of all teams, it is important to reach a shared understanding of each team's data retention needs prior to program end. As such, discuss and plan for data retention needs during program start-up to ensure the systems for collecting and storing program data will link smoothly.

Only retain what you need

Programs often gather a lot of data. For fear of deleting critical data, some program managers may choose to retain all data. Deleting data prematurely does pose an understandable risk (and this is covered more below). Yet, retaining data for too long – particularly participants' personally identifiable information (PII) – is equally risky.

The CaLP Principles and Operational Standards specifically recommend that you retain program participants' PII no longer than required:

Principle 7: *"Organizations should not hold beneficiary data for longer than is required unless they have clear, justifiable and documented reasons for doing so; otherwise, data held by the organization and any relevant third parties should be destroyed."*

Why get rid of data?

Retaining data longer than necessary increases the risk that information is leaked or accessed by "adversaries" (i.e., people whom you do not want to access your data). Additionally, as time goes by and staff inevitably change, it becomes more difficult to ensure that data is accurate, that new staff are aware of the limitations set on a certain data set and that a data source can be accurately traced.

An organization holding data is responsible for the data it holds – particularly program participant data. As such, it must be able to respond to requests from participants to access their personal data. The more data an organization holds, the more difficult this becomes, both in terms of the time required to answer a larger numbers of requests,

and the technical aspects of finding and isolating the relevant data. As a leading privacy expert warns: "If you have information, someone will demand you give it to them."

...but not too soon!

Discarding data too soon can also be harmful to programming. It is possible to discard data that partners or other stakeholders may need, and it can be very difficult or impossible to retrieve discarded data (if you have deleted it properly!). Depending upon your program type and funding structure, you may need to retain relevant data to justify decision-making and to refer back to for monitoring, evaluation and learning purposes.

WHAT HUMANITARIANS CAN DO:

Create a **Retention, Archival** and **Disposal (RAD) plan**

A RAD plan is a systemic way to manage program data. RAD plans can be customized for different situations. (See sample RAD plans links in the Additional Resources section at the end of this tip sheet.) In general, all RAD plans should describe:

- The **types of data** the organization must retain (and why)
- The **length of time** the data should be stored (and why)
- The **format** in which such data should be stored (and why)
- **How** the data will be retained, archived, transferred and destroyed
- **Who** is authorized to delete data (referred to as a "RAD Officer"), and who is responsible for confirming all organization data is properly destroyed before disposing of organization equipment. (Assigning these responsibilities to a job function (e.g., the M&E or IT officer) rather than a named team member can help ensure continuity of coverage if there are staff changes.)
- **Who** is covered by the plan and team member roles/responsibilities
- The **penalties** resulting from plan violations

All staff with a responsibilities related to this plan should sign that they understand the plan and pledge to uphold its tenets. Plans should also clearly state that no organization officer, employee or other representative is to modify, delete or destroy any data in violation of local, state, national, international or industry regulation.

Create an information inventory

A single program may contain various data sets and different approaches to gathering this data. The first step in managing data is to map out exactly what information is held, where it is held, and what the retention needs are for each type of data.

To do this, organizations often create an "information inventory" in conjunction with their RAD plan. An information inventory is a document that identifies steps within a program where major information groups are collected, and provides a complete list of the locations where they are held (e.g., on a certain computer, USB, hardcopy with the finance/program/M&E team in a field office or country head office, etc.).

Understand your legal environment

Different data retention laws apply depending upon the country/countries in which you work. Cross-border programs may face even more complicated regulations. If you are unsure which data retention laws apply to you – or are gathering particularly sensitive data – you may want to seek legal counsel to ensure you interpret and apply data retention laws appropriately.

Ensure good tagging and data provenance

Make sure the data you archive is understandable to someone unfamiliar with your program. This can mean labeling your files well, tagging information with labels that are clear to outsiders and/or leaving an inventory of the folders you have archived with a clear description of the folder structure.

Recording **data provenance** is the process of documenting data's origin and the path it took to end up in its current state (e.g., labeling how and when it was collected and how it has since been edited). Make sure to tag or label your data so someone can clearly tell where it came from, when it was collected, and how it was collected so that it can effectively be used in the future. As an example, when saving survey data, include the questions that were asked alongside it.

Be clear about what is shared with whom

Depending on your organization's policies, some staff may share files from personal accounts, rather than work ones. When these team members leave, their work accounts are disabled, but access to organizational files from their personal accounts often remains open. Wherever possible, keep work-related data on work-related accounts and encourage your team to do the same. If boundaries between work and personal accounts have already blurred, regularly verify who has access to what data and adjust sharing permissions accordingly. This might mean periodically removing access for personal email addresses that can log in to organization platforms or updating the authorized users of software tools each time new staff are hired.

Include a backup in your retention and archiving plans

If it is especially important to retain and archive certain data (e.g., in the event of a donor audit), make sure a backup of this data exists. Store critical data on a second hard drive in a different location or a second online location in the event the first version is lost or damaged. If you are storing a backup of PII, make sure that both the original version and the backup version are held with the same security and protection measures in place. Also, when it comes time to dispose of PII data in one location, the same procedure or actions are applied to the backup PII. Make sure the backup locations are clear (by listing in your information inventory, for example) so newcomers can keep track of critical data and the backup.

Plan for data usability

When deciding upon a format in which to save data, consider the length of time you will be saving it. Will someone with different software or operating systems still be able to open your data? As "dead formats" become more of a reality, make sure to save your data in common formats rather than "unusual" ones. The same principle applies to which cloud-based company you may choose; think how likely it is that the company will be in existence in a few years' time and opt for more established services, or open source options, rather than a new start-up, for long-term storage.

Clearly communicate your archival plan

If you will archive certain data, consider the length of time you will need to keep it and give clear instructions to anyone who may follow you. This might mean including a 'README' document along with the data on whatever storage device you choose, which states clearly when the data should be disposed of and how. Alternatively, make sure that a document (and maybe a calendar alert as well) with these instructions is evident to team members who may join after you (or others) have left.

Clearly communicate your archival plan

If you will archive certain data, consider the length of time you will need to keep it and give clear instructions to anyone who may follow you. This might mean including a 'README' document along with the data on whatever storage device you choose, which states clearly when the data should be disposed of and how. Alternatively, make sure that a document (and maybe a calendar alert as well), with these instructions is evident to team members who may join

after you (or others) have left.

Use secure data disposal tools

As the digital security guide *Security in a Box* states, "From a purely technical perspective, there is no such thing as a delete function on your computer." So, while you might think that emptying the "Recycle Bin" on your computer is sufficient, that document has not been deleted securely. When you are certain that you want to dispose of data, make sure you do so effectively. Refer to the open source tool *Eraser* and Security in a Box's *"Destroy Sensitive Information"* post for specific guidance.

ADDITIONAL RESOURCES:

10 things you should know about long-term data archiving. TechRepublic. July 2010.

Creating the records retention schedule you need. TAB. December 2012.

Deleting Personal Data. UK Information Commissioner's Office (ICO). An explanation of what is meant by "archiving", "deleting", "destruction" and "beyond use".

The importance of data retention policies. TechRepublic. July 2006.

Retaining Personal Data. UK Information Commissioner's Office (ICO). Principle 5 of the Data Protection principles: "you to retain personal data no longer than is necessary for the purpose you obtained it for." This is part of a useful data protection guide for organizations, but focuses on *personal* data.

Security in a Box: Destroy Sensitive Information. Frontline Defenders and Tactical Technology Collective. A how-to on destroying sensitive information.

Training curriculum on data retention and backup, with a focus on digital security. LevelUp.

Examples of RAD policies

- Central Kentucky Riding for Hope, INC *record retention and destruction policy.*
- Sheffield Health and Social Care *retention and disposal policy* (includes an implementation plan).

The Electronic Cash Transfer Learning Action Network is convened by Mercy Corps, with support from the MasterCard Center for Inclusive Growth.

MERCY CORPS

MasterCard | MasterCard Center for Inclusive Growth

elan
The Electronic Cash Transfer Learning Action Network