

DATA PROTECTION CASE STUDY: NIGERIA

Author: Norman Shamas

In March 2016, the Electronic Cash Transfer Learning Action Network (ELAN) brought together Mercy Corps' and Catholic Relief Services' Nigeria teams to discuss data protection in their electronic voucher programs and overall operations. Because data protection relies so heavily on context (e.g., internal policies on data sharing, national laws, and the parties interested in the data), the workshop focused on processes to create a culture of data protection and privacy. The two main processes discussed were **data flow mapping** and **risk assessments**.

WHAT CONSTITUTES PROGRAM DATA?

Program data includes all of the information collected and used throughout a program. In the programs discussed here, data includes beneficiary registration details, program details such as voucher value allocated, vendor details, voucher use (sales) data, operational planning and more.



DATA FLOWS ILLUSTRATED

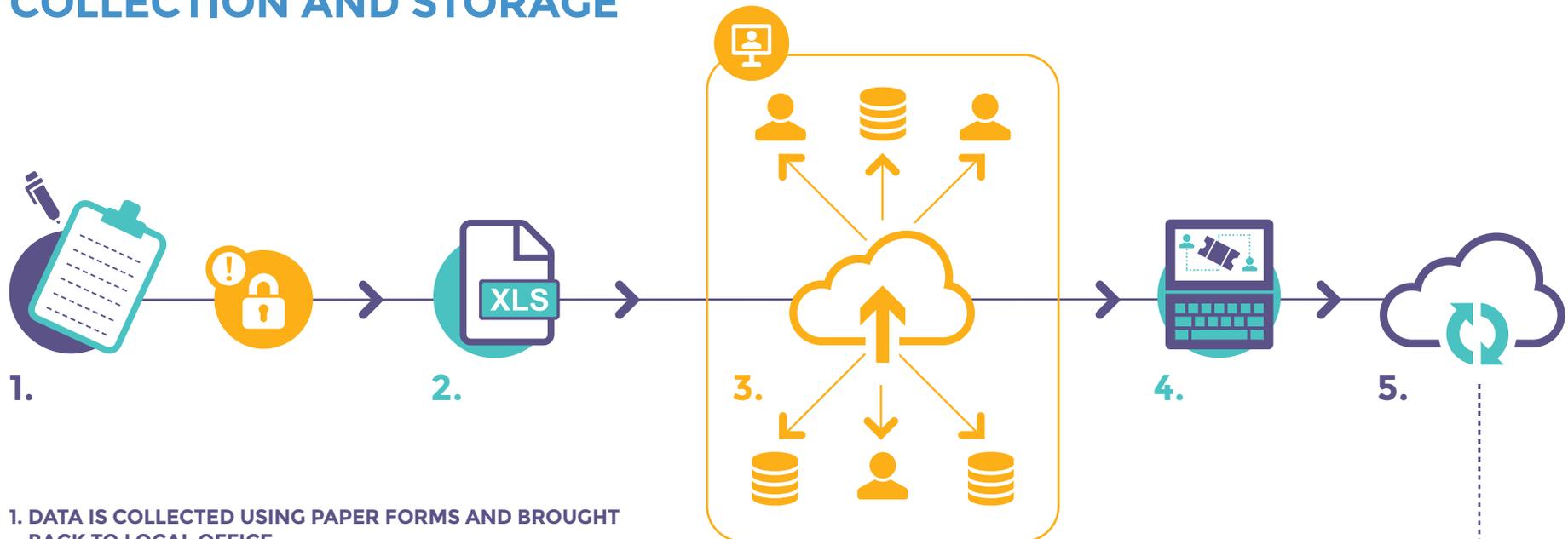
Each organization mapped its current data system to identify aspects that could be better secured or consolidated. These data flow maps served as the basis for the the rest of the workshop. Despite working with similar types of interventions and using the same platform to track food distributions, the organizations identified very different data flows and immediate risks based on analysis of their respective program implementation contexts.

PROCESS USED FOR DATA FLOW MAPPING

While there are a number of ways to map data flows, we used this process to guide the data flow mapping:

1. Identify tools and methods used to collect, store, sync, and/or backup data (e.g., cloud services, automated syncing with backup drives)
2. Arrange in a sequence that shows how data flows through the data lifecycle (i.e., from collection through deletion)
3. Go through the data flow again to identify more granular elements (e.g., identify the different documents stored in cloud service)
4. Add arrows between the methods and tools to illustrate the data flow. Use loops to show updates or cyclical data collection.

CASE 1: DIRECT IMPLEMENTATION USING MANUAL DATA COLLECTION AND STORAGE



1. DATA IS COLLECTED USING PAPER FORMS AND BROUGHT BACK TO LOCAL OFFICE

 *Securing data throughout the data lifecycle and in transit needs to be considered for both paper and digital data.*

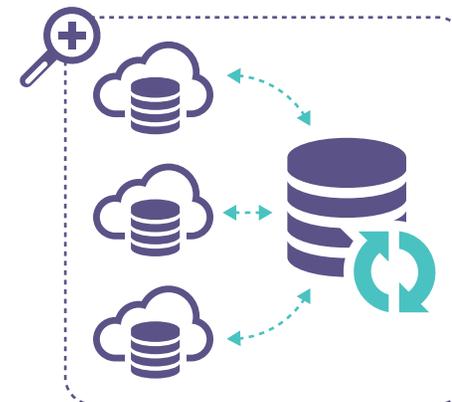
2. PAPER FORMS ARE TRANSCRIBED TO EXCEL FILES AS THE PRIMARY DATABASES

3. EXCEL DATABASES ARE UPLOADED TO A CLOUD FILE SHARING PLATFORM, WHICH SYNCs THE MOST UP TO DATE VERSION WITH EMPLOYEES AND A NUMBER OF DIFFERENT DATA SOURCES

 *In this scenario, the platform was an employee's personal account as opposed to a business managed account.*

4. DATA IS UPLOADED TO RED ROSE EVOUCHER PLATFORM

5. DATA FROM CLOUD FILE SHARING PLATFORM IS INCORPORATED INTO A DATA BACKUP PROCEDURES



Detail for step 5

CASE 2: REMOTE IMPLEMENTATION USING DIGITAL DATA COLLECTION AND STORAGE

1. DATA COLLECTION IS DONE DIGITALLY, WHETHER DIRECT BENEFICIARY DATA OR SALES INFORMATION

BENEFICIARY DATA

2.a. DATA IS DIRECTLY UPLOADED TO A CLOUD BASED DIGITAL FORMS PLATFORM

 Exporting data to an Excel file makes the data vulnerable since it is not longer subject to platform access and permission restrictions.

2.b. DATA IS SYNCED TO A LOCAL DEVICE, WHICH UPLOADS THE DATA TO ONE OF THE TWO CLOUD PLATFORMS

SALES DATA

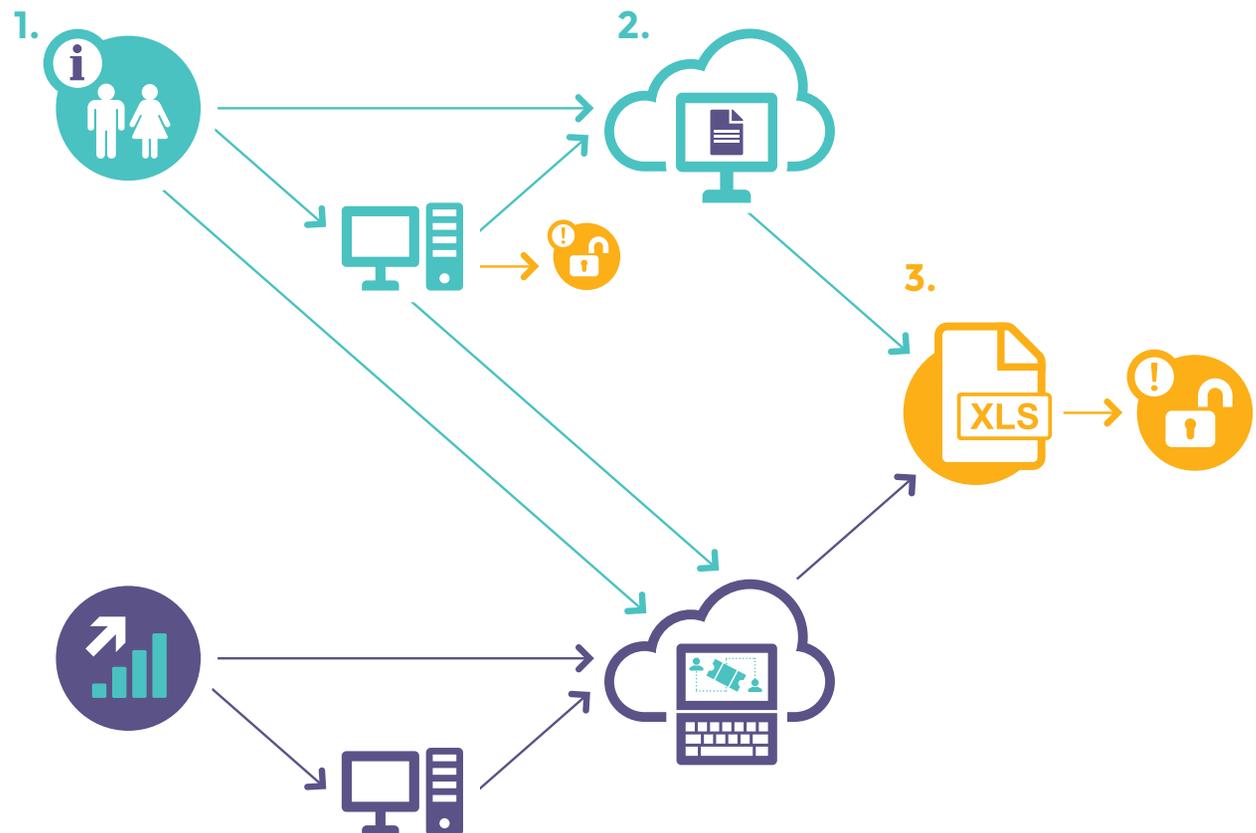
2.a. DATA IS DIRECTLY UPLOADED TO THE RED ROSE E-VOUCHER PLATFORM

2.b. DATA IS SYNCED TO A PROGRAM OFFICERS DEVICE, WHICH UPLOADS THE DATA TO RED ROSE E-VOUCHER PLATFORM

ONCE IN CLOUD PLATFORMS

3. DATA IS EXPORTED FROM THE CLOUD PLATFORMS TO AN EXCEL FILE FOR UPLOAD FROM THE DIGITAL FORMS SERVICE TO RED ROSE E-VOUCHER PLATFORM AND ANALYSIS ON LOCAL COMPUTERS

 Exporting data to an Excel file makes the data vulnerable since it is not longer subject to platform access and permission restrictions.



PROTECTING DATA IN TRANSIT

Protecting data in transit is one of the most overlooked aspects of (physical or digital) data protection. While the transmission method will change the specific security measures, the idea is the same: sensitive data requires trust in the transmission method and acknowledgement that some data may still be visible to everyone (i.e., metadata). Here are two tools to help protect data in transit:

HTTPS – The Internet was designed with collaboration rather than with security in mind. As a result, most traffic that is sent over the Internet (using HTTP) is visible to everyone. By using HTTPS, all the data is sent securely between the user and the website/service. Most browsers use a lock icon to easily show an HTTPS connection. HTTPS isn't available for all websites/services and it depends on the web service enabling this security feature.

Virtual Private Network (VPN) – If you are using an untrusted¹ network (e.g., public Wifi hotspot) and need to access services securely, a VPN is recommended. VPNs are setup by the user (not the service, like HTTPS) and protect all data that goes through the VPN. It is important to set up your own VPN or use a trusted VPN service: the VPN provider will be able to see your metadata and insecure traffic as if it were a normal Internet access point. Additionally, a VPN allows users to access internal networks and services and can create an additional layer of authentication for sensitive services.

NEW CHALLENGES WITH CLOUD SERVICES

Cloud services are an increasingly popular way to work effectively by continually sharing up-to-date data across different locations. One of the key advantages of using a cloud service is that the selected service

is responsible for managing the technical aspects of digital security, allowing your organization to focus on what you do best. However, cloud services raise several new challenges for organizations that use them:

- 1. Digital Security Literacy** – Humanitarian organizations now need to consider digital security in order to select cloud services and determine what and to what extent sensitive information should be stored using these services.
- 2. Interoperability** – When selecting cloud services, it is important to ensure that they are able to communicate with each other directly (interoperability). For example, if two cloud services are being used, one for data collection and one for monitoring, they are not interoperable if you need to download the data from the data collection service in order to upload into the monitoring system. Lack of interoperability can make it more difficult to maintain and protect sensitive data.
- 3. Cloud Service Management** – Centralized management of a cloud service by the organization should be a priority when selecting a service. When data is stored in a personal cloud account, the employee and organization take on a lot of legal liability if there is a data breach.
- 4. Data Terms of Use** – Terms of use define key concerns of data ownership and privacy requirements (e.g., service provider's rights to use stored data). If you are using a service to store sensitive data, having a separate data terms of use or clauses in a contract is necessary.
- 5. Legal Compliance** – Local laws about data sovereignty and data protection might limit the options and use of cloud services.

¹ An untrusted network is any network where you would be concerned about the network owner having access to information. This can be a "public" or "open" network or a network that requires login without a trusted party controlling the network (e.g., a hotel network).

THE BEST DATA PROTECTION IS NOT COLLECTING IT

Whenever data is collected and stored there is the possibility of that data being accessed by unauthorized users. Even though the best way to protect sensitive data is to not collect it, this is not always possible for humanitarian organizations. Instead, the focus should be on minimizing the data collected to what is essential. If personally identifiable information (PII) is not needed, do not collect it. At the beginning of the data lifecycle (i.e., the planning stage), consider how you can minimize data collected. For more tips on data minimization, see the [ELAN Data Minimization Tip Sheet](#).

Tip: If you plan to use PII only as a unique identifier, consider using a combination of other data to create a unique identifier or using an arbitrary identifier.

THINKING ABOUT RISK

Protecting data is a process of risk management based on the systems in place and local context. Risk assessment (or threat modeling) is a commonly used technique to understand the various risks to your valuable data, how to mitigate those risks, and which tools and processes to prioritize. A strong risk assessment considers more than just digital risks; at a minimum, it should also include physical (e.g., theft of devices), psychosocial (e.g., stress level and how that can affect following policies), and operational (e.g., trust among stakeholders) risks. Because risk can vary greatly based on context, a risk assessment can be modified to fit your specific system or environment.

A basic risk assessment can be conducted by asking a few questions (these questions can be modified for your specific system and/or environment).

- What data is valuable or can be damaging if it was not protected? What are some of the specific data points or objects that someone might be interested in?
- Who has an interest in this data? What are their estimated capabilities to access it?
- What happens if you lose control of the data or it is deleted?
- What measures can we take to protect or obfuscate the data?²

There is no single, correct way to run a risk assessment. During the workshop, the two organizations approached the risk assessment differently. One of the teams spent a lot of time identifying valuable data points and who is interested in them, while the other team focused on valuable files and devices and mitigation methods correlated to responsible parties internally. By taking these different approaches and reviewing the risk assessments together, both organizations were able to share findings for a more in-depth analysis.

As you do more risk assessments, you will get better at the process and identifying what needs to be protected, from whom, and potential mitigations. If possible, try and do risk assessments regularly with a diverse group of participants. Contextual knowledge (around technical and programmatic systems) is critical for useful risk assessments, so it is far better to bring knowledgeable team members together for this process than to have an information security expert independently conduct a programmatic risk assessment.³

² For details on obfuscation and how it differs from encryption, see “data obfuscation” in the glossary.

³ If you are creating your own technology platform or software, however, an information security professional's input throughout the development and deployment process is necessary.

Tips when thinking about what to protect and the consequences

- Data that is valuable or high risk might not be immediately obvious and might be a single data point within a larger data set.

Example: Dates and locations for distribution days

- If there are data or devices that have severe consequences and are easily accessible to an interested party, prioritize securing these.

Example: An unprotected spreadsheet that has beneficiary data has high consequences if not protected and anyone who can access the file (legitimately or illegitimately) can see all the data.

BREAKDOWN OF CAUSES OF DATA BREACHES



31%

Phishing, Hacking,
Malware



24%

Employee Action /
Mistake



17%

External
Theft



14%

Vendor



8%

Internal
Theft



6%

Lost or Improper
Disposal

Source: [Cyber Security Trend](#)

- Digital data might be most easily accessed through an unprotected device or human intervention.

Examples: Data stored on an unencrypted device can be accessed if someone has physical access to the device. People can be coerced (physically, legally, etc) into accessing the data or can neglect security policies when under high stress.

Tips when thinking about who could be interested in valuable data

- It is better to overestimate capabilities rather than underestimate.

Example: Criminal and terrorist organizations are increasingly recruiting technical experts and conducting cyber operations

- Sometimes groups are not as cohesive as they appear or there might be varying interests and capabilities from different parts of a single group.

Example: A government and its military might have different goals and capabilities to access valuable data

- Some actors can get protected data from different, more technical or data privileged organizations.

Example: Governments can often access mobile network data, including call records, voice recordings, and text messages.

THINKING ABOUT MITIGATIONS

There is no such thing as perfect security. Security is part of a specific context that needs to balance data protection and usability. If security measures make it too difficult for staff to complete their work, they will circumvent the security measures.

Sample Mitigations

Risk	Mitigation	Pros	Cons
Devices that contain valuable data can be physically accessed easily (e.g., stolen during a physical attack/raid or seized by authorities at a checkpoint)	Institute better device management policies to require strong lock codes/ passwords, prevent users from installing unauthorized apps, enforce full disk encryption to protect data on the device, and allow remote data wipes and tracking of devices.	<ul style="list-style-type: none"> ➤ Data on stolen or lost devices cannot be easily accessed without the lock code/password ➤ Users cannot install malicious application ➤ If a device is lost or stolen, the data can be deleted to prevent unauthorized access 	<ul style="list-style-type: none"> ➤ Need to define and enforce the policies, which might require additional responsibilities for staff or hiring of additional personnel ➤ If a lock code/password is forgotten, the data might be difficult or impossible to recover ➤ More devices required so that staff do not use personal devices for program purposes ➤ Users of the device will invest less care into a device they cannot customize and it can be treated with less care. In the case of the organization implementing remotely, locked down devices were provided to participating vendors
Beneficiary data is stored in an unprotected spreadsheet	Create a centralized database with more granular permission controls	<ul style="list-style-type: none"> ➤ Stronger permission controls can reduce unnecessary access to sensitive information while still providing access to the data points that are required for specific team members' work ➤ A centralized database prevents multiple versions of the data set 	<ul style="list-style-type: none"> ➤ Need to create and maintain the database and access portal, as well as permission levels and profiles, which requires time and resources

SOME GOOD PRACTICES THAT CAN BE EASILY IMPLEMENTED

The workshop concluded with providing some simple security practices that can be easily implemented and greatly decrease risk.

Passwords and authentication

Technical solutions for encryption and data protection are very strong and usually only protected by a password.

- Creating strong and unique passwords for all accounts and services is one of the best ways to protect your data. Unique passwords prevent someone from accessing multiple accounts and services if they figure out one password.
- Password managers make it possible to have a strong and unique password for every service without having to remember or write down all the passwords. They usually have a password generator as well!⁴
- When possible, use more than a password: two-factor authentication typically refers to a password (something you know) and a short-term, randomly generated code that is on a separate device (something you have). Many services, such as Google, provide support for two-factor authentication.

Secure Communications

People often need to work on or transmit sensitive data on networks that are not controlled by their organization or a trusted party.

- A Virtual Private Network (VPN) is a great way to protect data on untrusted networks, especially when using insecure (HTTP) web services.

- When using mobile networks (i.e., SMS and voice calls), assume that the mobile network operator and government (at a minimum) can access the content of your communications. If you need to discuss sensitive information through a phone call or SMS, you can provide some protection by obfuscating your conversation.

Note: *Obfuscation will only slow down an interested party from analyzing the communication and will typically require a secure channel of communication to agree on the obfuscation method(s). Since obfuscation can be easily defeated with enough data, it is also recommended that you change how you are obfuscating data frequently to mitigate an interest party collecting enough data to analyze your communications.*

- Use more secure communication tools when possible.

Example: *Skype is widely used, but its encryption methods have been broken and decryption tools are accessible online. If you need to share beneficiary data through screen sharing or a voice call, consider using a more secure alternative, such as: [Jitsi Meet](#), [talky](#), [Peerio](#), or [Wire](#).*

Incident Response

A commonly accepted belief in information security is that a breach is inevitable. So it is important to develop a digital security incident response plan. While this is not a new concept and physical security incident response is a common part of programs operating in conflict areas, the digital security aspect is usually not included. Like a physical incident response plan, appropriate roles and procedures to continue working after and learn from the incident should be defined.

When creating an incident response plan, it is important to think about roles and coordination between other parts of the organization. In the case that the security team does not oversee physical and

⁴ Some tools: [KeepassX](#) is a cross-platform tool that is free and open source; if you share passwords throughout the team, [1Password](#) has a great [team functionality](#).

digital security, it will be important to coordinate physical and digital security incident responses between the security teams.

BUILDING DATA PROTECTION INTO AN ORGANIZATION

As the importance of data protection is becoming more widely recognized, donors are starting to add required roles around data protection (e.g., DFID requires an information security plan, including an incident response process, to be developed and used for programs that it funds). Start thinking about roles and responsibilities for data protection at the planning stage.

Data protection is a process with a landscape that is changing quickly. Assessing risk does not happen once at the beginning of a program; instead it should be done continuously throughout. The organizations at the workshop determined that revising the risk assessment at least once a year would be manageable without adding too much of an additional burden on programming. Find a channel or forum to share security knowledge with other organizations that work in the same area to help others stay up to date on risks and the mitigations in the changing context.

ADDITIONAL RESOURCES

[ELAN Data Starter Kit](#)

[Responsible Development Data – Practitioner’s Guide](#)

[Responsible Data Handbook](#)

GLOSSARY

Breach/incident – A breach/incident is any event where data is illegitimately accessed or makes data unable to be legitimately accessed. This can come in many forms. Some examples are: illegal access to data from a 3rd party, malware (e.g., the growing threat of ransomware access to files), deletion of data, sharing sensitive information with a 3rd party, and denial of access to web services by a 3rd party attack.

Cloud file sharing/cloud service – A cloud service is any platform or service that is hosted and accessed entirely online with no data from the service being saved on a local computer. Cloud services can be: hosted by an organization’s infrastructure (i.e., self-deployed), created by an organization but hosted on a cloud hosting provider (e.g., Amazon Web Services), or created and deployed by a third party (e.g., Red Rose). One type of cloud service is cloud file sharing. This service allows individuals to collaborate and share the documents or other types of files while keeping the latest version available without needing to share via another medium. Common examples of cloud file sharing services are Microsoft’s OneDrive, DropBox, and Google Drive.

Data flow mapping – Data flow mapping (also known as network or data flow diagramming) is a technique to represent a system visually.

Data Obfuscation – Data obfuscation is a technique used to hide data in plain site by making it more difficult to analyze for the desired information. An example of data obfuscation is making it more difficult to use geographic data to identify where someone lives. In this case, data can be obfuscated by using a single location for all participants (e.g., community center) making it more difficult to identify where any particular participant’s household is. The primary difference between obfuscation and encryption is that encryption is meant to prevent access to data and obfuscation is meant to slow down analysis through some form of concealment.

Data Protection – Data protection refers to a number of concerns related to personal data privacy, collection and dissemination of information, and its legal and political considerations.

Metadata – Metadata is data about data/datasets; one of the earliest examples is card catalogs used in libraries. In the case of digital data, metadata typically refers to data that is automatically collected because of the protocol that is used to send the data. Metadata is public by design even when the data it describes is protected. For example, someone who is protecting their home address by using a P.O. box would be unable to prevent someone from seeing who is sending them mail to their P.O. box and what their P.O. box number is.

Personally Identifiable Information (PII) – PII is any information or grouping of information that can be used to identify a single person. Some easily recognizable examples of PII include: home address, passport number, and biometric data. There is also information that is not identifiable on its own, but can easily identify someone with additional information. Some examples are: IMSI/IMEI number, web cookies, and certain geodata.

Protected data – Protected data refers to any data that is not available to public. Data is typically protected because of privacy concerns or business needs.

Risk – Risk is anything that contributes an incident/breach. When conducting a risk assessment, it is important to also consider the possibility of the risk happening.

Threat – In the context of data protection, threat can refer to any method or any actor that can create a breach/incident. Threats can be further categorized into the medium the threat arises (e.g., threats to personnel in order to obtain information is a physical threat)

Two-factor authentication – Two-factor (or multi-factor) authentication is a technique in security to protect logins using more than a single piece of information. The two methods for authentication are usually: 1) something you know (e.g., password) and 2) something you have on you (e.g., one time password/code generated by another device). Two-factor authentication can also be used for services other than logins. For example, credit cards in Europe use RFID enabled cards and PINs as the methods of authentication for a purchase. In this example, a third-factor of authentication can be used by verifying someone's identity through an ID. While two-factor authentication greatly decreases the chance of attacks on authentication (e.g., accessing an account by having a user's password), it does not remove all risk.

Unique identifier (UID) – A unique identifier is a numeric or alphanumeric code that is unique to an entity within a given system. (e.g., a national ID number for an individual)

Virtual Private Network (VPN) – A VPN is a tool to network computers without requiring them to be located in the same physical space. In doing so, it allows a client computer to access servers and documents on an intranet.

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



MasterCard Center
for Inclusive Growth



Thanks to CRS and Mercy Corps staff in Nigeria and the US for their contributions to this case study.

