Changing the landscape of data and digital health solutions

COP21 Solutions

November 2020







Data.FI brings together leaders across the digital health and analytics landscape to harness the power of data to save lives

WHO WE ARE

Data.FI is a global health fieldsupport mechanism with a \$180M ceiling implemented 2019–2024 by:

- Palladium (prime)
- John Snow Inc.
- Right to Care
- macro-eyes
- IMC Worldwide
- Johns Hopkins University
- Cooper/Smith
- Jembi Health Systems

Data.FI is supported by a community of resources partners including BAO, Development Gateway, Fraym, IBM, Premise, Regenstrief, and others.



Digital health system enhancement and scale-up to transform health care



Data analytics that pinpoint health care inefficiencies and solutions



Decision support interventions to maximize the use of real-time data



Data standards and governance structures that optimize investments and ensure data quality

Data.FI supports USAID to:

- Reach those in need of HIV services
- Track clients
- Tailor client care
- Improve program quality
- Protect those we serve
- Maximize efficiencies
- Demonstrate accountability
- Ensure sustainability



Reaching those in need of HIV services

Data.FI Solutions

Population profiling and size estimation

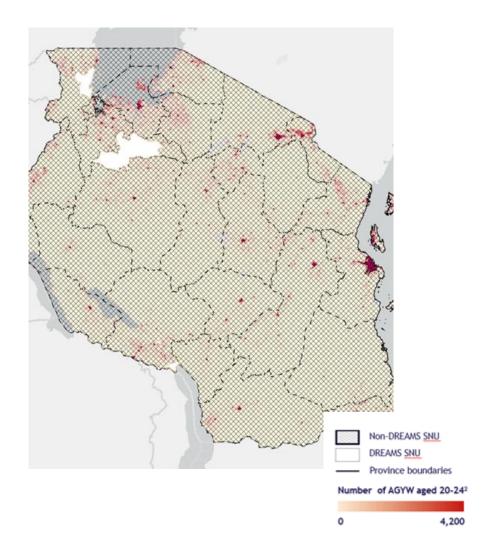
Population profiling and size estimation

Data.FI applies artificial intelligence/machine learning (AI/ML) techniques to generate subpopulation profile maps and size estimates to understand hyperlocal spatial distribution of populations at risk of acquiring HIV or treatment interruption.

We apply best practices from the literature to characterize risk, taking into account multiple overlapping vulnerabilities.

This method can be applied to:

- At-risk adolescent girls and young women (AGYW)
- Key populations (KPs)
- Orphans and vulnerable children (OVC)
- Pre-exposure prophylaxis (PrEP) candidates
- Voluntary medical male circumcision (VMMC) candidates



Data.FI is helping USAID reach AGYW in Uganda, Tanzania, Mozambique, Haiti, and eSwatini

We are supporting DREAMS programming by better estimating the density and location of AGYW considered at risk for HIV and estimating the size of those populations.

- Data.FI applied machine learning and spatial interpolation techniques to generate hyperlocal risk maps and size estimates of the population of AGYW
- Method uses population-based survey data and satellite imagery to create localized population information at a 1km resolution
- Risk profile maps visually represent the density and location of AGYW considered at risk for HIV
- In Uganda, these estimates were used to calculate saturation among current DREAMS districts, and identify high-priority districts for DREAMS expansion

Tracking clients

Data.FI Solutions

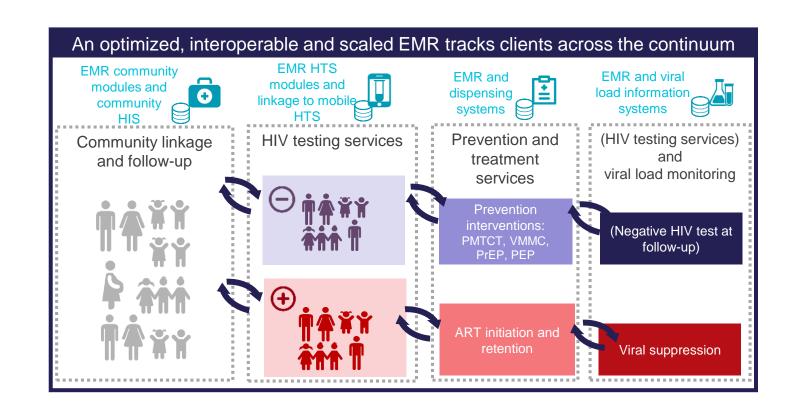
- Electronic medical record (EMR) systems
- OVC and DREAMS case management systems
- 3. Deduplication of client records

Optimizing and scaling EMRs

EMRs track clients across the continuum, generating data for improving clinical care, and ultimately informing client - centered approaches and enabling cohort analyses and program monitoring.

Data.Fl supports USAID and countries to optimize and scale EMRs by:

- Assessing existing EMRs for optimization and scale
- Aligning EMRs to treatment guidelines and reporting requirements from PEPFAR and countries
- Architecting, networking, and interoperating systems for exchange of information across community, testing, facility, dispensing, and lab systems
- Developing new modules for KP, OVC, PrEP
- Building in decision support for point of care EMR systems to facilitate quality care and differentiated care models
- Providing implementation and user support
- Facilitating linkages between community and facility systems through client scheduling and outreach for early missed appointments



Developing a global good: The LAMISPlus experience in Nigeria LAMISPlus v1.2 incorporations and provided in the Lamisplus v1.2 incorporation in the Lamisplus v1.2 incorporat

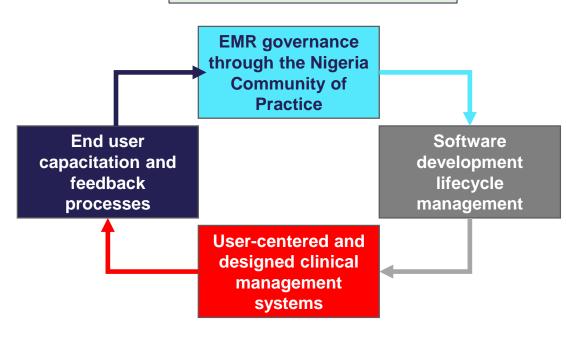
The LAMISPlus EMR system is an emergent global good application. It is a software tool that is free, open source, and used to manage, analyze, or transmit health-related data, with proven utility in several settings.

Data.FI established best-practices for the development of LAMISPlus including:

- Establishing a community of practice to ensure stakeholder ownership
- Using best-in-class software development and project management tools and methods, and publicly available technical documentation and source code for sustainability
- Applying responsive capacitation and user feedback modalities to ensure optimal software uptake and responsiveness to emerging user needs

LAMISPlus v1.2 incorporates these modules/functionalities:

- General clinic services
- HIV services
- COVID-19 Surveillance
- PMTCT services
- HTS services
- Report generation
- Administration module
- KP services



OVC case management systems

Customizable for your country

M&E Harmonization

Data.FI guides USAID and IPs in gathering requirements for adaptation of Global OVC Tracker Modules balancing reporting needs with data collection burden, with the aim of harmonizing systems, gaining efficiency, and increasing the reliability of the resulting data set.

Appropriate Technology

By leveraging the individual records module of DHIS2, called Tracker, we increase the ability to seamlessly push data into DATIM; mobile devices appropriate for community data collection allow decentralized data collection.

Performance Monitoring

Access to real-time OVC MIS supported by USAID ensures standardization of data collection and indicator calculation between partners and facilitates performance monitoring of case management before the end of the reporting period, improving our collective ability to meet targets.

Sustainable

Once developed and rolled out to IPs, Data.FI builds the capacity of government or local IPs to maintain the OVC Tracker modules into the future; USAID reduces overall maintenance costs by investing in one centralized system.

Data.FI is streamlining OVC, DREAMS, and sexual violence prevention data management in Zimbabwe

COUNTRY CONTEXT:

6 IPs, each with its own system for collecting and reporting data, including 80

different data capture forms

→ inefficient and costly

Data are shared with USAID monthly

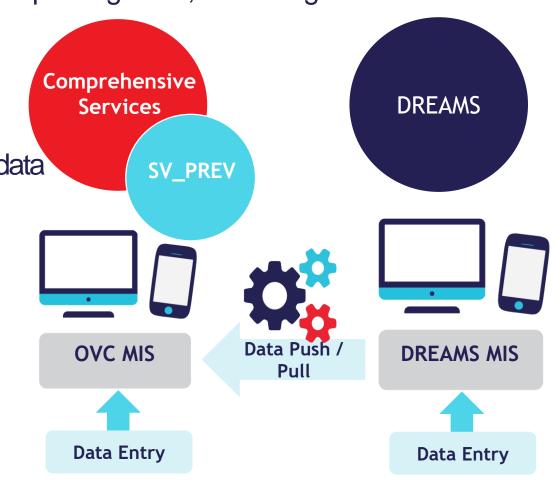
→ no data for real-time decision support

Partners manually calculate, aggregate, transfer data

→ data quality and reliability issues

IMPACT

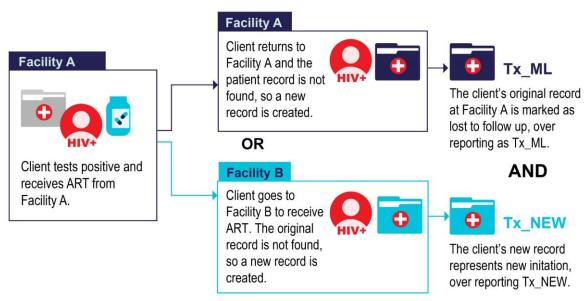
- A harmonized set of 6 data capture tools
- DHIS2 Tracker module that captures individual-level data and calculates indicators when multiple events are recorded and datestamped for one individual



Deduplicating client records

To improve data quality and client care



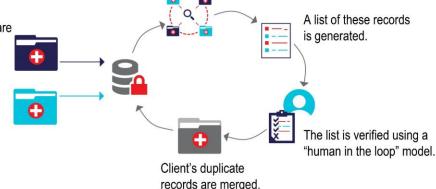


In the **Red Scenario**, Facility A over-reported TX_NEW and TX_ML_LTFU; in the **Blue Scenario**, Facility A over-reported TX_ML_LTFU while Facility B over-reported TX_NEW

THE SOLUTION

The deduplication algorithm identifies that the two records correspond to one unique client using text similarity or distance metrics.

Client's duplicate records are entered into a centralized data warehouse at the above-site level.



At a systems level, deduplication proceeds through the stages in the graphic on the right: where records are pooled, an algorithm searches for likely duplicates, there is a measure of human verification (in some form, and not in perpetuity), and duplicates are linked.

A patient-matching model for South Africa

In South Africa, Data.FI supported the National Department of Health to develop and deploy a supervised machine learning-based patient-matching model for patient records.

The patient-matching model matches over 20 million patient records within the national TB/HIV system and establishes the mechanism for other client records ingested within the InfoHub to be matched and unified.

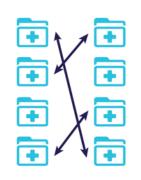
The matching model will allow the SA HIV program to:

- Identify clients who have become lost to follow-up at a facility but present themselves elsewhere for treatment, or tests captured through another data source
- Trace clients who have become lost to follow-up at one facility and present themselves at another facility, flagging them as having transferred out of one facility and transferred into another facility
- Augment existing test results for a client by overlaying the client record with lab results, filling in the blanks for clients whose test result record appears incomplete in the record

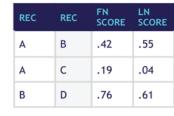
DEDUPLICATION

Anne-Marie	annemarie
+44 507 532	44507532
N'gethe	ngethe
NYC	newyorkcity

1. Preprocessing: Clean and standardize records



2. Indexing: Identifying candidate record pairs



Comparing: Quantify similarity between variables



4. Classifying:
Determine if
records are linked



Tailoring client care

Data.FI Solutions

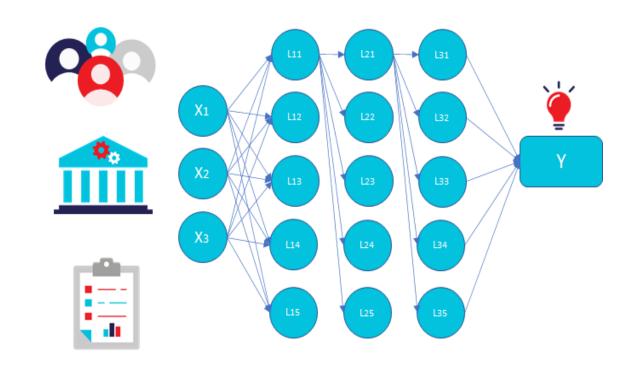
Predictive analytics and client phenotyping

Predictive analytics

ML/AI models identify patterns in data to predict which clients will face various outcomes across the clinical cascade and therefore to inform client-centered care.

With the proliferation of non-traditional data, rapidly advancing algorithms, and ever-expanding computational power, we have the tools to get more from our data, tailor client care, and improve clinical outcomes.

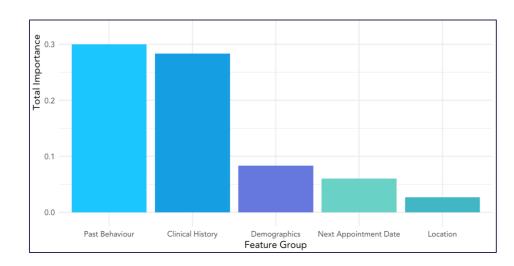
Our method can be used to predict which individuals are most likely to test positive, face co-infections (TB, COVID), experience treatment interruptions, be virally suppressed, or respond to different interventions.



Data.FI is predicting which clients will face treatment interruption in Mozambique and Nigeria

In Mozambique, the model identified the 20% of clients at greatest risk of treatment interruption. Of these, 75% interrupted treatment. IPs can use predictions to flag high-risk clients proactively and direct support services prior to visit dates, or to identify which clients already late for visits are most likely to self-return.

Scaling the Solution: This process can be replicated in any country with client-line data, with minimal input from IPs. The model can be integrated into information systems to automatically and continuously identify clients at high risk of treatment interruption.



Client behavior, such as history of lateness for appointments, and clinical history, such as time on treatment, were the most important factors in making predictions.

Demographic and locational factors had less predictive power.

Improving program quality

Data.FI Solutions

- 1. Epidemic Control Rooms
- 2. Digitized viral load tracking

Epidemic Control Rooms

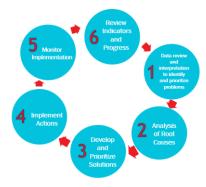
HIV Epidemic Control Rooms allow decision makers to analyze data in real time for continuous program improvement supported by change management practices.

Technology Enabled



HIV Epidemic Control Rooms using technology-based data analytics and visualization platforms integrate and/or triangulate data from multiple sources (e.g., EMRs, lab, and pharmacy records).

Standardized Methodology



Standardized methods for actionable data review processes include root cause analysis, asking the right questions to better interpret data, monitoring actions, feedback across levels, and capturing lessons learned.

Rapid Course Correction



Weekly data review allows service providers, community service managers, and decision makers from IPs, OUs, and government to closely monitor HIV data in real time and compare performance across sites and regions.

Data.FI is improving HIV clinical outcomes across the cascade through epidemic control rooms in Nigeria

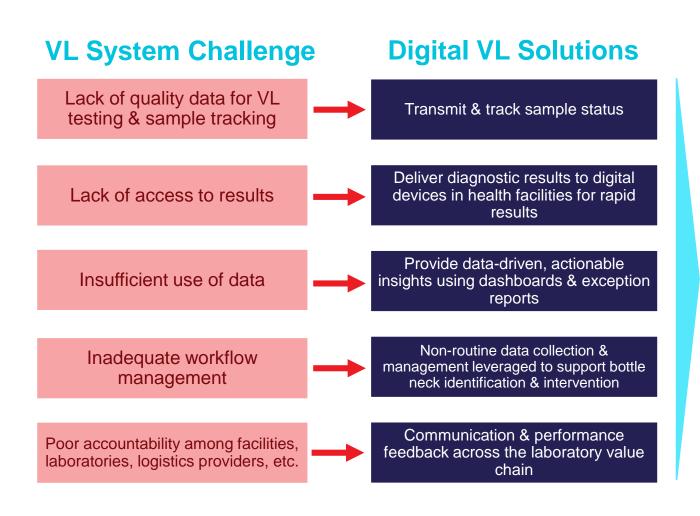
Prevention: Post ECR intervention, the average percent of PrEP initiation increased from 13% to 112% of the annual target.

Retention: Our analysis revealed that 44 facilities had program losses, with a total of 11,564 persons assumed to be LTFU. Post ECR intervention, 84% (9,703) of clients previously LTFU were tracked back to care.

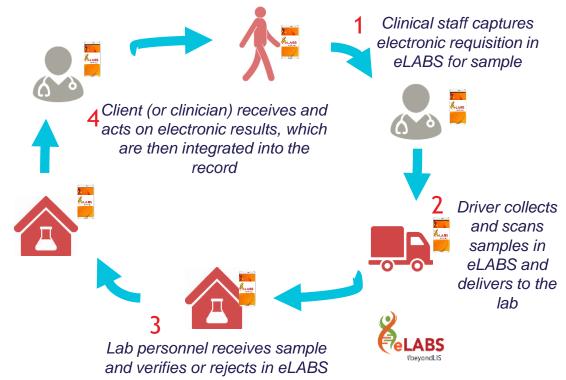
Viral Load Testing: Post ECR intervention, the average percent of eligible patients with documented VL testing results increased from 44% to 70%. In the seven lowest performing facilities, the average increased from zero to 79%.



Digitized viral load tracking



eLABS is currently used in **over 1,000 facilities** in **Zambia** and **South Africa**, with facilities reporting a **60% improvement in total turnaround time** for viral load results.



Protecting those we serve

Data.FI Solutions

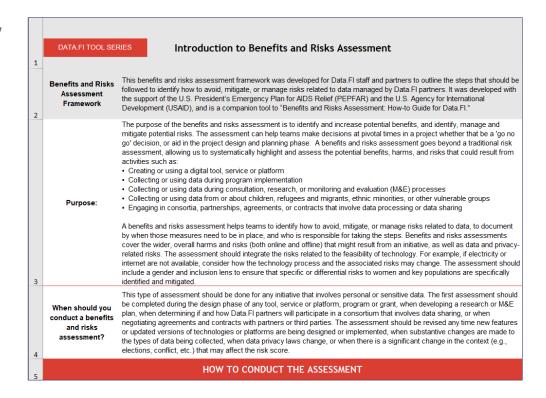
Data protection support

Data protection support

With the emphasis on collecting, integrating, and using client-level data, Data.FI supports USAID and partners to securely manage client data and develop technologies that protect client data.

Data.FI supports USAID and partners to:

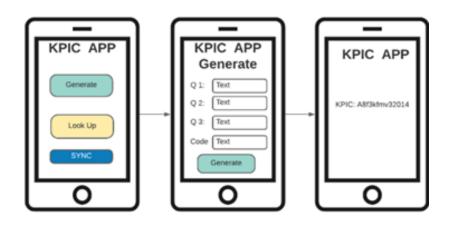
- Conduct benefits and risks assessments of biomarkers and other data-capture initiatives
- Develop data management plans that protect client data, with a focus on local partners
- Design technologies that protect client data, such as unique IDs based on privacy by design principles

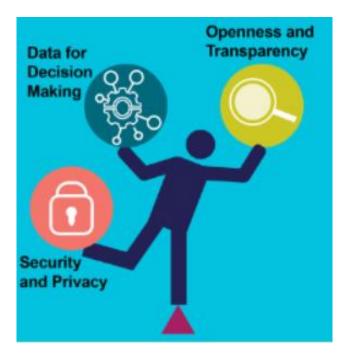




Data.FI is protecting client data in Uganda and Burundi

In **Uganda**, Data.FI is using a "privacy by design" unique identification approach in developing a mobile and web-based application for key populations to access HIV services safely. The alphanumeric key population identifier code (KPIC) permits health facilities to securely access clients' individual health records, without requiring the client to disclose their identity. This unique ID can be used without an expensive security infrastructure.





In **Burundi**, Data.FI conducted a biometrics benefits and risk assessment as a unique ID solution. Data.FI is now working with USAID and the government to document and implement a risk mitigation strategy.

Maximizing efficiencies

Data.FI Solutions

- 1. Data warehouses that support advanced analytics
- Efficiency analysis and costing

Centralized data warehouses

A central data warehouse matches client records and unifies data across disparate systems, providing comprehensive data for program management, predictive analytics, and cohort analyses.

Centralized Warehouse

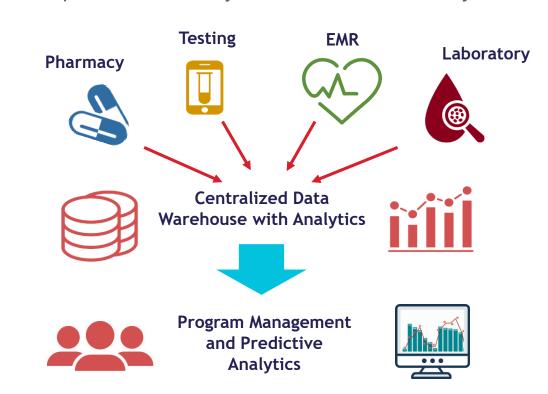
- Consensus and standards building for interoperability
- Ingestion of data sources from across the HIV continuum
- Data security protocols aligned with country regulations

Customizable Data Visualizations and Reports

- Real-time program analytics and access to high-frequency data
- Customizable and interactive dashboards for use in Epidemic Control Rooms
- Pre-programmed PEPFAR/MOH reports

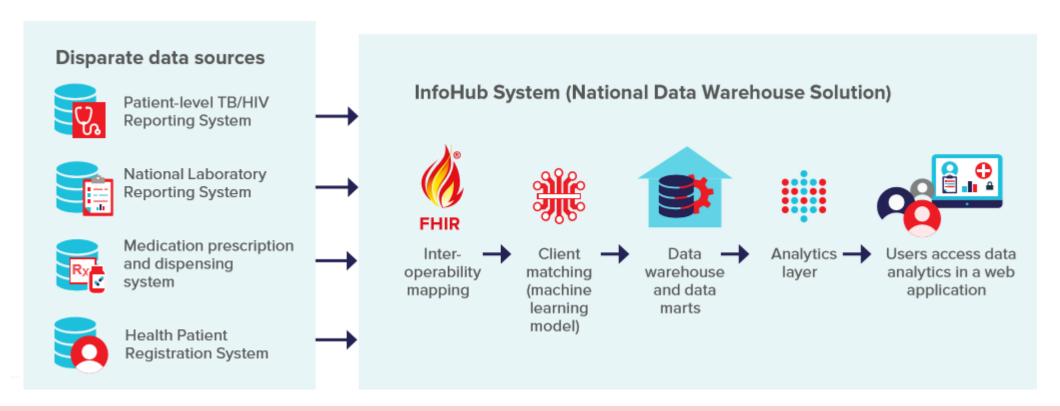
Surveillance and Predictive Analytics

- Client matching and deduplication across the HIV cohort
- Cohort analyses for assessing longitudinal outcomes and true LTFU
- Case-based surveillance



InfoHub in South Africa: An integrated analytics platform

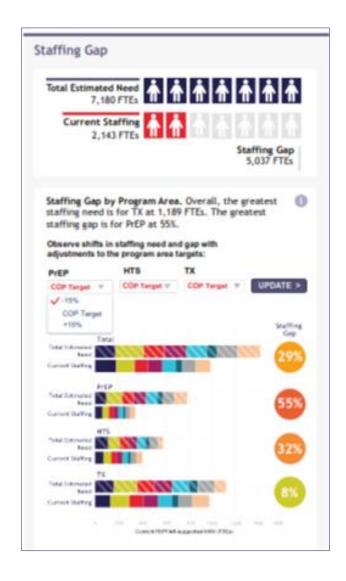
In South Africa, Data.FI is supporting the National Department of Health to build a centralized data warehouse that integrates disparate HIS data into insightful analytics, designed primarily by and for key policy and program decision makers, data analysts, data scientists, and epidemiologists.



The InfoHub uses a machine learning algorithm that matches patients across the datasets, in order to minimize duplicate records. The InfoHub is built using open-source, non-proprietary digital technologies with Metabase, used to display analytics outputs designed for various user cadres and data use cases across the HIV continuum.

Efficiency analysis and costing

- Human resources for health (HRH) optimization. Data.FI supports USAID to optimize HRH for maximal health impact through models such as our HRH planning tool.
- Value-based care. Data.FI works with USAID and partners to incorporate client value markers into existing continuous quality improvement interventions.
- **Efficiency analysis and costing.** Data.FI analyzes the efficiency of different service delivery models and provides cost estimates for services and systems.
- Modelling essential supplies. Data.FI supports USAID to forecast medical supplies and equipment needs, as well as facility requirements (expansion/repurposing) at country and subnational levels.



Demonstrating accountability

Data.FI Solutions

- 1. Data quality improvement
 - 1. Data quality composite scores
 - 2. Anomaly detection
- 2. Real-time, rapid data collection

Data Quality Composite Score Tool (DQS)

The Data Quality Composite Score Tool allows decision makers to automatically evaluate the quality of data and identify areas for improvement

Standardized Method



Provides data quality scores (DQS) in three dimensions: completeness (number of submitted vs expected records), coherence (numerator being less or equal than denominator), and consistency (presence of outliers).

Wide Application



Designed to work with data from any other data collection tool or software and is intended for use by staff above the facility level to assess data quality before expensive investments in audit processes in the field.

Data Quality Improvement



Results allow decision makers rapidly assess data quality dimensions across IPs, regions or sites, identify and address data quality gaps as well as compare performance over different time periods—including post-intervention.

Data.FI is improving the quality of HFR data through use of the DQS tool in Nigeria and West Africa

Nigeria and West Africa countries (Burkina Faso, Ghana, Liberia, Mali, Senegal, and Togo): For 12–37 weeks of HFR data reviewed, all detected data quality issues were resolved by the partners, resulting in 100 percent end line scores.

Moving forward, countries expressed interest in using the DQS tool to review and correct data at regular intervals.



Photo by Robert Glod, courtesy of Flickr







Data Quality Composite Score Tool

he Data Quality Composite Score Tool is a toolkit designed by Data.FI to automatically evaluate the data quality score f HIV data from organizational units supported by the President's Emergency Plan for AIDS Relief (PEPFAR). The tool is ovides data quality scores (DQS) in three dimensions: completeness (measures the number of ter than its denominator), and consistency (measures the number of outliers). The tool has ork with data from any other data collection tool or software, as long as data are prepared and

tor is available in the dataset; a measure of the

Improving Data Quality of High-Frequency Reports (HFR)

routine stakeholder meetings to review and



Impance data by Data FF, quarierly data quality tredback meetings, and national-level blue

o measure the quality of HFR data and changes in data quality over time. Data FI of Assume the quanty or new one and changes in one quanty over one, used it developes a data quanty of socials some that measures quanty of reported data using the following eight HFR indicators: HTS_TST and reported data using the following eight HFR indicators: HTS_TST

variation rules: HTS_TSTP*MTS_TST_POS: TX_NEWP*TLD_NEW; TX_CURRO*TLD_CURR; TX_PVLS_D>=TX_PVLS_N. Each of the four validation rule checks had a 25% weight in the control of the four validation rule checks had a 25% weight in the control of the four validation rule checks had a 25% weight in the control of the four validation rules checks had a 25% weight in the control of the four validation rules checks had a 25% weight in the control of the four validation rules checks had a 25% weight in the control of the four validation rules checks had a 25% weight in the control of the four validation rules checks had a 25% weight in the control of the four validation rules checks had a 25% weight in the control of the four validation rules checks had a 25% weight in the control of the four validation rules checks had a 25% weight in the control of the four validation rules checks had a 25% weight in the control of the four validation rules checks had a 25% weight in the control of the four validation rules checks had a 25% weight in the control of the four validation rules checks had a 25% weight in the control of the four validation rules checks had a 25% weight in the control of the four validation rules checks had a 25% weight in the control of the four validation rules checks had a 25% weight in the control of TX_PVLS_Qo=TX_PVLS_N: each or the loss valuation has checks had a 2018 weight in the cone score calculation; thus, when all health facilities under a particular iP met a validation rule, a 25% of the cone of th





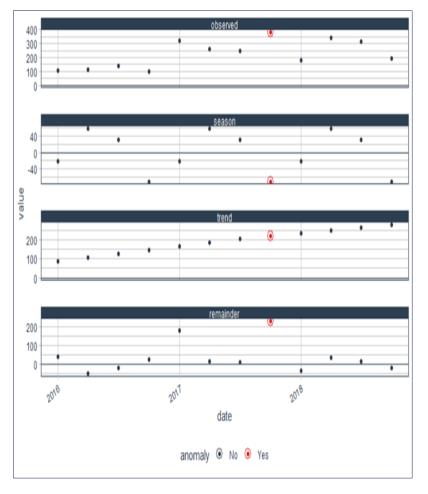


Anomaly detection

We apply algorithms to rapidly and efficiently flag anomalous data points in large datasets to identify likely data quality and performance issues.

Anomaly detection algorithms identify data points that deviate from normal values. They can help to flag atypical facility-level performance or patient outcomes. These approaches learn directly from the data provided. They do not require historical data, nor do they require any assumptions about where anomalies are likely to occur.

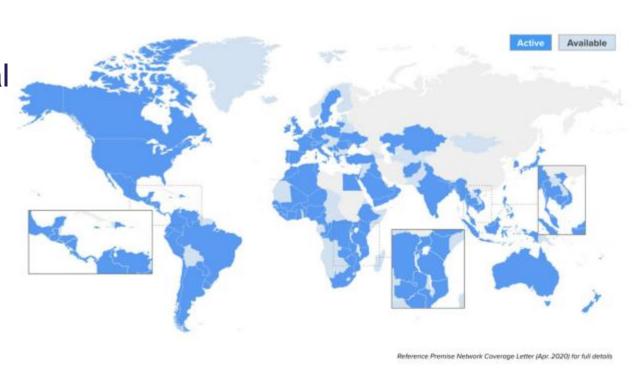
Facility	TX_CURR	TX_NEW	TX_PVLS	TX_ML	TX_RTT	PMTCT_ART	TB_ART	TB_STAT	HTS_TST
Alpha	59	4	55	4	4	NA	NA	NA	34
Beta	55	NA	47	3	5	NA	NA	NA	34
Gamma	577	52	456	143	162	25	1	2	344
Delta	64	3	57	5	6	4	NA	NA	24
Epsilon	49	NA	46	NA	4	NA	NA	NA	36



Real-time, rapid data collection

Data.FI can crowdsource insight in real time.

Data.FI partner PREMISE has the largest global crowdsourcing network in the world. PREMISE's data contributors can meet critical on-the-ground data needs within 24–48 hours, capturing both population- and site-level data. PREMISE uses a cloudbased platform to assign data capture tasks, manage, analyze, and visualize data for users.



Ensuring sustainability of investments

Data.FI Solutions

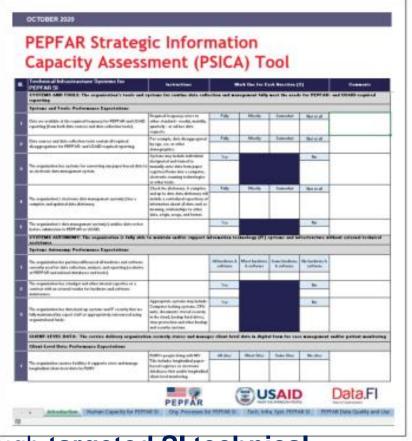
- PEPFAR Strategic Information
 (SI) Capacity Assessment
- 2. SI curriculum development and local partner training
- 3. HIS coordination and governance

Local partner capacity strengthening

Using our **PEPFAR Strategic Information Capacity Assessment (PSICA) tool,** Data.FI facilitates:

- Implementation of the online survey
- Analysis of results
- Joint mission-partner review and action planning





Data.FI supports missions and LPs to address capacity gaps through targeted SI technical assistance, including online training and virtual coaching.

Data.FI built an online training course for LPs in HIV Treatment Cascade Quality Assurance.

Data.FI is building LP capacity in Haiti, Kenya, Malawi, Tanzania, Zimbabwe, and the DRC

Data.FI implemented the PSICA tool with eight LPs.

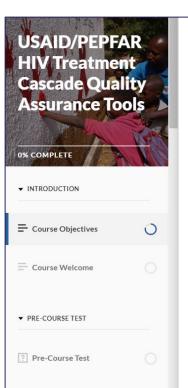
The process has been really helpful. The way [the questions] are crafted they are quite easy to follow and easy to understand.

-Grace Kumwenda, Chief of Party, IHDC/Malawi

The structure of the questions was really good because it enabled us to look at the *M&E* component in a holistic way.

-Victor Kanje, M&E Specialist, USAID/Malawi

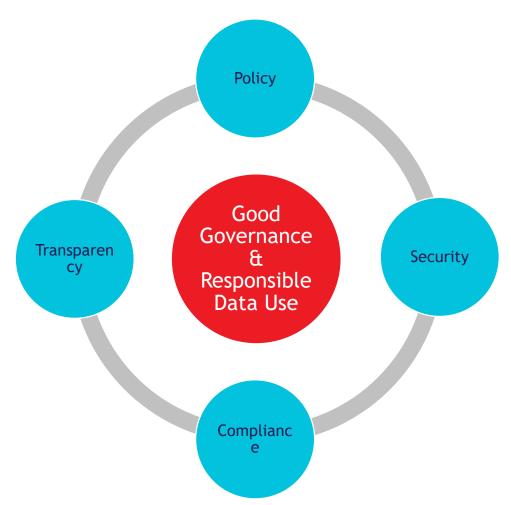
To address data quality issues in reporting and routine quality improvement efforts, Data.FI built an online training course for LPs and trained 25 LP staff in HIV Treatment Cascade Quality Assurance.



Welcome to the USAID/PEPFAR Data Quality Assessment Tools Online Course! This training focuses on the use of USAID/PEPFAR's Data Quality Assessment (DQA) tools and aims to provide learners with guidance and tips on the implementations and applications of DQA to improve HIV reporting. In this course, participants will be introduced to 3 new DQA tools that will ensure the optimization of data quality and use.



Strengthening HIS coordination and governance



Data.FI strengthens governments and partners to coordinate and sustain digital health investments by:

- Building consensus among stakeholders on common standards for system alignment
- Co-developing information system enhancements across government and partners with robust communities of practice
- Strengthening interoperability across disparate systems
- Supporting governments' eHealth strategies

Data.FI is strengthening local HIS leadership in Nigeria and Burundi

We bring stakeholders together to capitalize on local expertise and ensure sustainability and country ownership.

In **Nigeria**, through the LAMISPlus EMR community of practice, we brought together informatics staff among seven partner organizations. We held two **informatics bootcamps**, providing individal **mentoring** and learning sessions for local partners to collaboratively build open-source data modules into the national EMR.

We put in place **best practices in software development** including sharing programming code across
partners in GitHub and use project management software
platform JIRA to assign and monitor progress.

In **Burundi**, we supported the creation of the government-led SIDAInfo technical working group, which is collaboratively building a web-based version of the SIDAInfo EMR and supporting the development of unique client ID.



USAID/Nigeria 's IT manager, Toks Aluko, confers with the Data.FI
Nigeria team during the informatics bootcamp held in Lagos, Nigeria.



From left to right: Ekoche Ikeh (NACA), Bababunmi (FHI 360), and Ngodoo (Heartland Alliance) discuss the data analytics and mapping of indicators for different LAMISPlus modules. Photo by Data.FI.

Mechanism Information

PAD Number	936-8000
DUNS (Palladium)	186295239
Mechanism Name (FactsInfo)	Data for Implementation (Data.FI)
Mechanism ID (FactsInfo)	82099
Award Number (Cooperative Agreement)	7200AA19CA00004
Prime Partner	Palladium International, LLC
Sub Partners	JSI, JHU, Right to Care, Cooper/Smith, IMC Worldwide, Jembi Health Systems, Macro-Eyes
Start Date	April 15, 2019
End Date	April 14, 2024
Total Estimated Cost (+90% ceiling available)	\$179,895,772.99
Agreement Officer (AO)	Adrienne Shade (ashade@usaid.gov)
Agreement Officer Representative (AOR)	Emily Harris (emharris@usaid.gov)
Please Copy All Correspondence To	Data.Fl@usaid.gov
Project Director	Jenifer Chapman (jennifer.chapman@thepalladiumgroup.com)







FOR MORE INFORMATION

Emily Harris, Data.FI AOR, USAID Office of HIV/AIDS emharris@usaid.gov

Jenifer Chapman, Data.FI Project Director jenifer.chapman@thepalladiumgroup.com

Data for Implementation (Data.FI) is a five-year cooperative agreement funded by the U.S. President's Emergency Plan for AIDS Relief through the U.S. Agency for International Development under Agreement No. 7200AA19CA0004, beginning April 15, 2019. It is implemented by Palladium, in partnership with JSI Research & Training Institute (JSI), Johns Hopkins University (JHU) Department of Epidemiology, Right to Care (RTC), Cooper/Smith, IMC Worldwide, Jembi Health Systems and Macro-Eyes, and supported by expert local resource partners.

This presentation was produced for review by the U.S. President's Emergency Plan for AIDS Relief through the United States Agency for International Development. It was prepared by Data for Implementation. The information provided in this presentation is not official U.S. government information and does not necessarily reflect the views or positions of the U.S. President's Emergency Plan for AIDS Relief, U.S. Agency for International Development or the United States Government.