

PEPFAR Data Governance Guidance

April 2021



U.S. DEPARTMENT OF STATE

Contents

- 1. Introduction 3
- 2. Policy Directive and Background 3
- 3. PEPFAR Data Governance Structure 3
- 4. PEPFAR Data 4
 - 4.1 PEPFAR data streams 4
 - 4.2 PEPFAR data systems 6
- 5. Data Access 7
 - 5.1 Improving access to data for HIV epidemic control..... 7
 - 5.2 Open data exceptions 7
 - 5.3 Data generated at foreign military sites..... 8
 - 5.4 Data on key populations 9
 - 5.5 Data on Peace Corps volunteers 9
 - 5.6 Non-PEPFAR-supported data 9
 - 5.7 Comingled data 9
 - 5.8 Data access platforms 9
 - 5.9 Internal USG access..... 9
 - 5.10 External access 10
 - 5.11 External access process 11
 - 5.12 Data access roles and responsibilities 11
- 6 Data Security 12
- 7 Data Quality and Standards 13
 - 7.1 Data formats 14
 - 7.2 Objectivity and utility 14
 - 7.3 Integrity..... 14
 - 7.4 Requests for corrections 14
 - 7.5 Data quality roles and responsibilities 15
 - 7.6 Documentation and metadata 15

1. Introduction

Data, and the information derived from data, are assets for PEPFAR, its partners, academic and scientific communities, and the public at large. This document establishes general policy for managing data generated through PEPFAR-funded programming (hereinafter “PEPFAR data”). The PEPFAR program leverages data to improve our ability to achieve HIV epidemic control and increase transparency and accountability. The utility of data used for decision-making hinges on the degree to which data meet established quality criteria and are securely organized and accessible.

This document complements relevant policy directives, guidance, and authorities, including the Federal Data Strategy.¹ PEPFAR is managed and overseen by the U.S. Department of State's Office of the U.S. Global AIDS Coordinator and Health Diplomacy, which implements data policies consistent with any applicable guidance, including guidance of State Department's Bureau of Information Resource Management. This PEPFAR data governance guidance is also informed by data management policies of the PEPFAR implementing agencies: U.S. Agency for International Development (USAID); Department of Defense (DoD); Department of Commerce (DoC); Department of Labor (DoL); Department of Health and Human Services (HHS); Peace Corps and; Department of the Treasury. This PEPFAR data governance guidance document outlines roles and responsibilities for critical data management functions that support the PEPFAR program.

2. Policy Directive and Background

White House Executive Order 13642, OMB Circular A-130, and the OMB Open Data policy M-13-13 set relevant policy for the management of government information.^{2, 3, 4} According to OMB Circular A-130, “federal information” refers to information “created, collected, processed, maintained, disseminated, disclosed, or disposed of, by or for the Federal Government in any medium or form.”

3. PEPFAR Data Governance Structure

The U.S. Department of State's U.S. Global AIDS Coordinator & U.S. Special Representative for Global Health Diplomacy issues this PEPFAR data governance guidance based on consultation with PEPFAR implementing agencies.

This PEPFAR data governance guidance is implemented by data stewards. The term “data steward” refers to a role assigned to PEPFAR staff under their existing position to ensure adherence to data management guidance. PEPFAR data stewardship encompasses a set of roles and responsibilities that a variety of individuals undertake in relation to data at different points in data lifecycles.⁵ PEPFAR data are

¹ Federal Data Strategy, <https://strategy.data.gov/>

² White House Executive Order 13642, Making Open and Machine Readable the New Default for Government Information. <http://www.gpo.gov/fdsys/pkg/FR-2013-05-14/pdf/2013-11533.pdf>

³ Office of Management and Budget. OMB Circular A-130, Revised. Management of Federal Information Resources, <https://www.whitehouse.gov/sites/whitehouse.gov/files/omb/circulars/A130/a130revised.pdf>

⁴ Executive Office of the President. OMB Memorandum M-13-13 Open Data Policy-Managing Information as an Asset. <https://www.whitehouse.gov/sites/whitehouse.gov/files/omb/memoranda/2013/m-13-13.pdf>

⁵ STATE/IRM/BMP/OCA. 5 Fam 630, Data Management Policy, (CT:IM-274; 05-26-2020). <https://fam.state.gov/FAM/05FAM/05FAM0630.html>

organized into data streams needed to execute the management of PEPFAR programs. Each PEPFAR data stream has its own lifecycle and is subject to PEPFAR data governance guidance and data stream-specific supplementary guidance that is issued by Primary Data Stewards, which are a designated team or working group. Primary Data Stewards are responsible for identifying authoritative versions of data, clearing data for public release, resolving data quality issues, and authoring and maintaining metadata.

Questions about PEPFAR Data Governance Guidance should first go to the Primary Data Steward. S/GAC's Office of Program Results and Impact Monitoring for Epidemic Control (PRIME) convenes an Interagency Data Systems Committee consisting of representatives from the implementing agencies. The committee is chaired by the Deputy Coordinator for PRIME at S/GAC. The committee chair advises the U.S. Global AIDS Coordinator. The committee may provide strategic direction and guidance for PEPFAR data collection efforts.

4. PEPFAR Data

The focus of this document is on the management and distribution of PEPFAR data which, for purposes of this guidance, are data sets, reports, studies, and dashboards generated with PEPFAR support rather than information prepared for the administration, management, and operations of PEPFAR. This policy does not apply to press releases or similar communications.

4.1 PEPFAR data streams

Listed below are the primary PEPFAR-supported data streams, sources of data stream-specific documentation, and the respective Primary Data Stewards. Additional PEPFAR guidance may further document data governance considerations, such as data approval processes, access, and roles and responsibilities pertaining to a specific data stream.

- **Monitoring, Evaluation, and Reporting Indicators (MER):** PEPFAR program indicators for targets and results.
 - Primary Data Steward: S/GAC PRIME-DUIT
 - Documentation: PEPFAR Monitoring, Evaluation, and Reporting Indicator Reference Guide⁶; and Data Quality Assurance protocols
- **Site Improvement Through Monitoring System (SIMS):** quality management data from standardized tools that assess adherence to PEPFAR standards of HIV-related service delivery at the site and above site levels.
 - Primary Data Steward: S/GAC Program Quality
 - Documentation: SIMS Implementation Guide, Site Assessment Tool, Above Site Assessment Tool
- **Non-service delivery and above site investments:** unstructured program data that are not tracked by MER. They focus on the health system rather than direct health care service provision and are not allocated to specific patients or persons. These activities benefit all health system users and may result in support for health system functioning, increased effectiveness, improved efficiency, and/or greater equity.

⁶ PEPFAR Guidance. <https://www.state.gov/reports-pepfar/>

- Primary Data Steward: S/GAC PRIME-DUIT
- Documentation: PEPFAR Country Operational Plan Guidance for all PEPFAR Countries
- **Expenditure Reporting:** the amount of funds spent by PEPFAR implementing partners via Country/Regional Operational Plans.
 - Primary Data Steward: S/GAC PRIME-PET
 - Documentation: Expenditure Analysis of PEPFAR Programs Guidance, PEPFAR Financial Classification Guidance
- **Budget:** initial planned funding levels approved each year in the PEPFAR Country Operational Plans (COP) and Regional Operational Plans (ROP), Headquarters Operation Plans (HOP), Technical Leadership and Support (TLS), Administration, Working Capital Fund, and other budget plans. A COP/ROP is the vehicle for documenting U.S. government annual investments and program activities in HIV/AIDS for each PEPFAR operating unit (country or region). Funding may be comprised of new appropriated funding and applied pipeline from multiple fiscal years. Adjustments to the budget are made as conditions warrant throughout the year.
 - Primary Data Steward(s): S/GAC PRIME-PET, S/GAC Management and Budget
 - Documentation: Planning Level Letters, Approval Memos, FACTS Info NextGen Reports, Budget Trackers, Reports from the Global Financial Management System (GFMS)⁷
- **Outlay:** All payments by the federal government that track back to appropriations accounts for implementing mechanisms per agency.
 - Primary Data Steward: S/GAC Management and Budget
 - Documentation: Obligation & Outlay Reports, End of Fiscal Year Report, Reports from GFMS, Transfer Memos
- **Organization unit (OU) hierarchy:** geographic relationships, in tabular form, that define the arrangement of administrative units and health facilities in DATIM, which is an instance of DHIS2.⁸
 - Primary Data Steward: S/GAC PRIME
- **Spatial Data:** points and polygons that capture the geometry of administrative units and health facilities.
 - Primary Data Steward: S/GAC PRIME
 - Documentation: PEPFAR Spatial Data Lifecycle
- **Sustainability Index and Dashboard (SID):** a tool completed periodically by PEPFAR teams and partner stakeholders to sharpen the understanding of each country's sustainability landscape and to assist PEPFAR and others in making informed HIV/AIDS investment decisions.
 - Primary Data Steward: S/GAC Financial and Programmatic Sustainability
- **Partner Work Plans:**
 - **Budget template**
 - Primary Data Steward: S/GAC PRIME-PET
 - Documentation: S/GAC-specific guidance governs this information stream.
 - **Program narrative**

⁷ PEPFAR Guidance. <https://www.state.gov/reports-pepfar/>

⁸ DHIS2 Documentation. Organisation units. <https://docs.dhis2.org/en/home.html>

- Primary Data Steward: Agency-specific
 - Documentation: Agency-specific guidance governs this information stream.
- **Budget narrative**
 - Primary Data Steward: Agency-specific
 - Documentation: Agency-specific guidance governs this information stream.
- **Surveys and Surveillance:** data on epidemic features such as HIV prevalence; national HIV incidence; pediatric HIV prevalence; HIV viral load suppression; behavioral prevalence in the general population; behavioral prevalence among high-risk populations; and transmission of drug resistant HIV strains.
 - Primary Data Steward: Agency-specific
 - Documentation: Agency-specific guidance governs this data stream.
- **Implementation Science:** research that informs delivery and scale-up of efficacious interventions to improve HIV prevention, care, and treatment
 - Primary Data Steward: Agency-specific
 - Documentation: Agency-specific guidance governs this data stream.
- **Evaluations:** the systematic collection and analysis of information about the characteristics, outcomes, and impact of programs and projects
 - Primary Data Steward: Agency-specific
 - Documentation: Evaluation Standards of Practice⁹
- **Commodities:** information on health commodities, prices, and delivery destinations.
 - Primary Data Steward: Agency-specific with country team coordination
 - Documentation: Agency-specific guidance governs this data stream

4.2 PEPFAR data systems

Most PEPFAR-supported data stream lifecycles originate in transactional systems such as PEPFAR's health information system, Data for Accountability, Transparency and Impact (DATIM). These systems are optimized for data collection and data entry. The PEPFAR Data Hub (PDH) is the data warehouse, or repository, for data collected by transactional systems across the PEPFAR program. PDH consolidates, stores, and packages PEPFAR-supported data for internal analysis and review as well as public dissemination.

Panorama is a web-based business intelligence tool that allows U.S. Government users to visualize quarterly data. Panorama Spotlight is a public-facing version of these data. PEPFAR SharePoint supports knowledge management across the Implementing Agencies, both at headquarters and operating unit levels. FACTS Info NextGen is not owned by PEPFAR but is the system of record for foreign assistance planning, budgeting, and reporting, including PEPFAR assistance.

⁹ PEPFAR Guidance. https://datim.zendesk.com/hc/en-us/article_attachments/360040023292/PEPFAR_evaluation_standards_of_practice_v3.1_October_2019.pdf

5. Data Access

5.1 Improving access to data for HIV epidemic control

This PEPFAR data guidance seeks to make PEPFAR data accessible, to the extent consistent with applicable laws and regulations, to maintain transparency and improve our ability to achieve HIV epidemic control. This data access policy is intended to enable the PEPFAR community to take full advantage of PEPFAR data and responsibly expand public access to valuable information. PEPFAR encourages its partners, the academic and scientific communities, and the public at large to make broad use of PEPFAR data for innovative scientific, technological, analytical, and other applications when those data are shared responsibly and in accordance with this policy.

A key feature of data accessibility is security level. PEPFAR uses the OMB data categorization that describes data as public, restricted public, or non-public.¹⁰

- Public: Data asset is or could be made publicly available to all without restrictions.
- Restricted Public: Data asset is available under certain use restrictions (e.g. made available to select implementing partners, multilateral organizations, or researchers under certain conditions).
- Non-public: Data asset is not publicly available and is only available for internal use by the Federal Government.

Primary Data Stewards determine the security level of data under their responsibility and clear the public release of data, such as through Panorama Spotlight, Data.gov, or ForeignAssistance.gov.

5.2 Open data exceptions

Certain sensitive data may warrant, or applicable law or regulation may require, withholding certain data from public or other disclosure. For example, privacy, confidentiality, and other considerations related to some PEPFAR data – such as patient-level data, data concerning foreign militaries, or data concerning the exact locations where vulnerable populations congregate – may be inappropriate for public dissemination. The handling of sensitive but unclassified information generated within the PEPFAR program may be subject to guidelines established by relevant implementing agencies.^{11, 12, 13}

Consistent with the appropriate treatment of sensitive data, as well as with applicable laws and regulations, relevant PEPFAR implementing agencies should not publish any information that may provide an unfair competitive advantage in procurement. For example, budget data, to the extent it is

¹⁰ Federal Enterprise Data Resources. DCAT-US Schema v1.1 (Project Open Data Metadata Schema).

<https://resources.data.gov/schemas/dcat-us/v1.1/>

¹¹ State/DS/SI/IS. CT:DS-284. "12 FAM 540 Sensitive but Unclassified Information (SBU)",

<https://fam.state.gov/fam/12fam/12fam0540.html>

¹² HHS/OCIO/OIS. Document #: HHS-OCIO-OIS-2019-05-004, Version #: 2.2 "HHS Rules of Behavior for the Use of HHS Information and IT Resources Policy",

<https://www.hhs.gov/web/governance/digital-strategy/it-policy-archive/hhs-rules-of-behavior-for-the-use-of-hhs-information-and-it-resources-policy.html>

¹³ USAID/M/CIO/IA. File Name: 545_061819. "ADS Chapter 545. Information Systems Security."

<https://www.usaid.gov/sites/default/files/documents/1868/545.pdf>

made public, should be redacted for procurement-sensitive funding information for awards that have not yet been made.

TBD Mechanisms for awards that have not yet been made by the awarding agency will not be shown as separate mechanisms in any public data sets. These mechanisms will be aggregated until awarded so as to avoid revealing any procurement sensitive information related to the expected amount of the award. Once awarded, the mechanism will be both named and its amount will be identified and reported in budget data sets.

To accommodate policy or legal considerations, PEPFAR data streams are made available with different levels of identification and granularity for U.S. government use and the public. PEPFAR attenuates risks associated with potentially sensitive data through geographic aggregation where appropriate.¹⁴

The most granular data used for routine planning and monitoring are at the site level, which is either a clinical site or a low-level administrative unit (e.g. a district). Data may be anonymized to allow publication at fine levels of granularity. Additionally, the program may use cell suppression for small values to reduce risks associated with disclosure of these data.

5.3 Data generated at foreign military sites

A portion of PEPFAR data relates to foreign military sites, such as bases, barracks, or military hospitals. PEPFAR support may be USG DoD military-to-partner military (mil-mil), USG (e.g. CDC or USAID) civilian-to-partner military (civ-mil), or USG DoD military-to-civilian (mil-civ) activities. These categories describe physical locations where PEPFAR supports services. They do not describe funding agencies or mechanisms.

- Military-military (mil-mil) are DoD programs with a primary focus to serve military populations at military site locations. Beneficiaries may include civilians who are provided with services at military sites (e.g. a military base, barracks, or other installation).
- Civilian-military (civ-mil) are usually CDC or USAID programs serving military personnel at military sites. Beneficiaries may include civilians who are provided with services at military sites.
- Military-civilian (mil-civ) programs primarily focus on serving civilian populations at civilian sites.

PEPFAR records data on PEPFAR mil-civ activities at the level of granularity specified by the data stream with no special exception. All mil-mil and civ-mil data, at any level of aggregation, are non-public.

In DATIM, data originating at military sites are aggregated to a designated military unit, which is a national-level aggregation. Implementing partners should enter all mil-mil and civ-mil data to the military unit in DATIM irrespective of the funding agency. Prior to adding a site to DATIM, DATIM Site Administrators must verify that sites are not military sites (i.e. mil-mil or civ-mil). Military sites may, through error, appear in DATIM. When these sites are discovered, data reported against the sites must

¹⁴ National Institute of Standards and Technology. Guide to Protecting the Confidentiality of Personally Identifiable Information (PII). Special Publication (NIST SP) - 800-122, <https://www.nist.gov/publications/guide-protecting-confidentiality-personally-identifiable-information-pii>

be rolled up to the military unit. Following data roll up, the military site must be deleted in a manner consistent with applicable laws and regulations.

5.4 Data on key populations

Key populations are especially vulnerable to stigma, harassment, abuse, and discrimination and sometimes subject to criminalization. These groups include sex workers, gay men and other men who have sex with men, people who inject drugs, transgender people and prisoners. The PEPFAR policy is that all PEPFAR data and PEPFAR-funded publications related to key populations are subject to heightened attention to privacy, confidentiality, and security. PEPFAR does not make facility-level data on key populations publicly available.

Site names occasionally indicate a type of key population outreach or beneficiary served at the site. The DATIM site list is reviewed semiannually for these cases. When a site is tagged as sensitive, its name is masked within PEPFAR's analytic data systems (e.g. Panorama).

5.5 Data on Peace Corps volunteers

PEPFAR does not make public data that can link Peace Corps volunteers to a specific site or subnational geographic area that may identify the presence of a volunteer.

5.6 Non-PEPFAR-supported data

PEPFAR systems may store data owned by entities outside of PEPFAR. These are data that were not generated through PEPFAR funds, such as data shared with PEPFAR but owned by a partner government or private entity. Examples could include commercial data, results imported from Ministry of Health data alignment, and non-public spatial data created through a private foundation. Unless copyright license or data use agreement explicitly permits further distribution by PEPFAR, these data are non-public in their most granular form and at any level of aggregation.

5.7 Comingled data

Combinations of data from different sources take on the most restrictive category of the constituent data elements. For example, an analysis that leverages MER data from Spotlight, public domain CC0 1.0 license data, and facility-level SIMS data may not be publicly releasable since the SIMS data may contain sensitive information related to key populations. Similarly, analytic products derived from combinations of data, for example PEPFAR-supported and partner government data, should be cleared by all owners prior to public release. To support awareness of data lineage, data stewards should clearly document data sources for all data sets, reports, dashboards, and analytic products.

5.8 Data access platforms

Data management tools and platforms used in the PEPFAR program include Panorama, DATIM, FACTS Info, PEPFAR SharePoint, ForeignAssistance.gov, ArcGIS Online, and others. PEPFAR implements open data policy through Panorama Spotlight, State.gov/PEPFAR, Data.gov, and ForeignAssistance.gov.

5.9 Internal USG access

Internal USG access to data is defined by roles. As data move through the approval process within a given data stream, the number of people with access to the data tends to increase, but always in

relation to a defined stewardship role, such as implementing partner data clerk, GIS Specialist, DATIM Site Administrator, Strategic Information Liaison, PEPFAR Program Manager, or Expenditure Analysis Advisor.

Implementing agencies administer user rights to shared S/GAC-managed systems in accordance with memorandum of understanding and enforce role-based access to data in support of PEPFAR programmatic goals. Implementing agencies may leverage these shared data in agency-specific internal systems provided they maintain the same role-based access.

5.10 External access

Panorama Spotlight is the primary platform used by PEPFAR for making PEPFAR data available to the public.¹⁵ Unless otherwise noted, data that appear on the public dashboard are in the public domain through the CC0 1.0 Universal Public Domain Dedication.¹⁶ All public PEPFAR-supported data sets should have a public metadata record on Data.gov. PEPFAR updates public dashboards with cleared data on a routine schedule, much of which is governed by the PEPFAR Data Calendar.¹⁷ Decisions to release new indicators or levels of detail to existing indicators to the public should be documented, presented to the Interagency Data Systems Committee, and disseminated to Agency Deputy Principals for their feedback.

Table 1 describes for the primary data streams hosted on Panorama Spotlight the approximate level of data granularity associated with each access level.

Table 1. Default access categories and the primary PEPFAR data streams on Spotlight

Table 1.

Data stream/ Access level	Non-public	Restricted public	Public
MER (quarterly)	<ul style="list-style-type: none"> Named facility-level data for most sensitive indicator subsets MER narratives 	<ul style="list-style-type: none"> Named facility-level data for indicator subsets 	<ul style="list-style-type: none"> Targets and results by subnational unit and named mechanism for most indicators Masked facility-level data for indicator subsets
SIMS (quarterly)	<ul style="list-style-type: none"> Named site-level, agency-level, and IM-level data Responses to questions below SIMS final scores by Core Essential Element 		<ul style="list-style-type: none"> SIMS final scores by Core Essential Element at subnational unit 1 (SNU1)-level
ER (annually)			<ul style="list-style-type: none"> Expenditure by mechanism
Budget (annually)	<ul style="list-style-type: none"> TBD mechanisms 		<ul style="list-style-type: none"> Budget by mechanism

¹⁵ PEPFAR Dashboards, <https://data.pepfar.gov>

¹⁶ CC0 1.0 Universal Public Domain Dedication, <http://creativecommons.org/publicdomain/zero/1.0>

¹⁷ PEPFAR Data Calendar, <https://datim.zendesk.com/hc/en-us>

5.11 External access process

Public data published to Panorama Spotlight are downloadable in tabular form. General requests for restricted public data for which S/GAC is the Primary Data Steward can be made by filling out a request form available on <https://data.pepfar.gov/faqs>.

Requests for OU-specific restricted public data will be directed to OU teams. OU teams have discretion over how country-specific restricted public data may be disseminated. PEPFAR Country Coordinators may, in consultation with interagency stakeholders, share country-specific restricted public data among, for example, civil society, implementing partners, researchers, or partner governments.

S/GAC will lead responses on requests for multi-OU restricted public data. OU teams will be given an opportunity to provide a justification for the rejection of a request in accordance with exceptions to open data. S/GAC and the implementing agencies will track and document restricted public data requests and communicate restricted public data dissemination to the Interagency Data Systems Committee.

Requests for restricted public data

	OU-specific data	Multi-OU data request
S/GAC role	Sends to OU level	Leads response
OU role	Leads response	Can veto

All individuals and entities receiving restricted public data sets should sign a PEPFAR Data Use Agreement (DUA), which sets the terms for the use of restricted public data. External access to agency-specific restricted public data created with PEPFAR support, such as from evaluations, surveillance and surveys, and implementation science, may be requested from the Primary Data Stewards of the respective source.

5.12 Data access roles and responsibilities

5.12.1 Interagency Data Systems Committee

A standing Interagency Data Systems Committee advises the committee chair on data access. This group can recommend the release of data and considers possible justifications for data release exemptions. With the development of new data streams and data management scenarios, the Interagency Data Systems Committee supports the creation additional guidance that advances the responsible sharing and use of PEPFAR data.

5.12.2 S/GAC's Primary Data Stewards

S/GAC's PRIME Team convenes the Interagency Data Systems Committee. S/GAC's PRIME Team is also responsible for maintaining systems that assure role-based access to most PEPFAR data. These systems, tools, repositories, and business intelligence platforms include The PEPFAR Data Hub, Panorama Spotlight, Panorama, DATIM, ArcGIS Online, and others. PRIME coordinates among all S/GAC Primary Data Stewards in an external data access process which establishes data security categorization for PEPFAR-supported data streams. Finally, PRIME works to see that PEPFAR data that has been cleared for public release by responsible Primary Data Stewards, reside on open platforms and have recorded metadata on Data.gov.

5.12.3 Agencies

Implementing agencies should ensure that PEPFAR-supported contracts, grants, cooperative agreements, and other instruments include existing agency-specific data access-related guidance and that recipients fulfill those requirements. Where applicable (e.g. SIMS data) agencies provide, in machine readable format, complete and approved data according to the calendar for each data stream.

5.12.4 Country teams

PEPFAR Country Teams play a role in handling requests for OU-specific restricted public data. They also provide complete, approved data according to the calendar for each data stream. Country Teams confirm that mil-mil, civ-mil and key population data are appropriately reported. Country Teams also monitor site lists for explicit reference to key populations (e.g. Sex Worker Friendly Clinic #8), work with implementing partners to identify official site names, and update DATIM accordingly.

5.12.5 Implementing partners/grantees

PEPFAR-funded recipients provide quality data in machine readable formats according to the applicable reporting calendar and data definitions for each relevant data stream.

6 Data Security

Data security policy seeks to ensure appropriate access to data through the implementation of risk-based security requirements.¹⁸ Such security policy addresses physical, electronic, and procedural aspects of protecting information. Security policy aims to protect data from inadvertent or malicious inappropriate disclosure and non-availability of data due to system failure and user error.

PEPFAR adheres to standards for categorizing information and information system security according to a range of risk levels. PEPFAR also adheres to information security requirements (i.e., management,

¹⁸ UNAIDS. 2006. Guidelines on Protecting the Confidentiality and Security of HIV Information. http://data.unaids.org/pub/manual/2007/confidentiality_security_interim_guidelines_15may2007_en.pdf

operational, and technical security controls) for information and information systems in each such category.^{19, 20, 21, 22}

Data stream-specific documentation traces the chain of custody from data generation to the data's state of rest and should address the following data security components:

- Responsibilities
- Data collection and use
- Data sharing and release
- Physical security
- Electronic data security
- Audit and monitoring
- Retention²³

Role based access control is implemented for each PEPFAR system to ensure users have appropriate access privileges for each system and the data stream residing on that system. For example, obtaining a user account for DATIM and Panorama require a role that justifies access. An Interagency System Security Forum (ISSF) reviews PEPFAR technology as it pertains to security and privacy implementations, technology configuration, data security standards, policy, and procedures. The ISSF meets monthly or as needed.

7 Data Quality and Standards

A data driven epidemic response requires quality data. PEPFAR is committed to the collection and dissemination of accurate, reliable, and unbiased information in a clear, complete, and unbiased manner. PEPFAR follows the OMB Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by Federal Agencies.²⁴ OMB Guidelines set forth basic information quality standards and note that the more important the information, the higher the standard to which it should be held. PEPFAR weighs the costs and benefits of information quality in all data streams and is committed to creating and managing data that have objectivity, utility, and integrity.

PEPFAR data that are influential scientific, financial, or statistical information should be reproducible according to commonly accepted scientific, financial, or statistical standards. Such PEPFAR data that are

¹⁹ National Institute of Standards and Technology. Standards for Security Categorization of Federal Information and Information Systems (Federal Information Processing Standard [FIPS] 199). <http://csrc.nist.gov/publications/fips/fips199/FIPS-PUB-199-final.pdf>

²⁰ National Institute of Standards and Technology. Special Publication SP 800-60 Vol. I Rev. 1, Guide for Mapping Types of Information and Information Systems to Security Categories (August 2008). <http://dx.doi.org/10.6028/NIST.SP.800-60v1r1>

²¹ National Institute of Standards and Technology. Special Publication SP 800-60 Vol. II Rev. 1, Appendices to Guide for Mapping Types of Information and Information Systems to Security Categories (August 2008). <http://dx.doi.org/10.6028/NIST.SP.800-60v2r1>

²² National Institute of Standards and Technology. Special Publication SP 800-37 Rev. 2, Risk Management Framework for Information Systems and Organizations: A System Life Cycle Approach for Security and Privacy (December 2018). <https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-37r2.pdf>

²³ Executive Office of the President. OMB Memorandum M-12-18 Managing Government Records Directive. <https://www.archives.gov/records-mgmt/m-12-18.pdf>

²⁴ OMB, Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by Federal Agencies, <https://www.gpo.gov/fdsys/pkg/FR-2002-02-22/pdf/R2-59.pdf>

for internal use only undergo rigorous, documented robust checks. The Primary Data Steward is responsible for ensuring that data quality checks are performed and provides approval following any necessary data transformation activities. Quality standards for PEPFAR data appear in data stream-specific documentation.

7.1 Data formats

Open data formats support the widest array of potential users. In keeping with open data principles, PEPFAR publishes data sets in machine readable, non-proprietary formats (e.g. comma-separated values (csv), shapefile (shp), etc.) for all data streams.^{25, 26}

7.2 Objectivity and utility

In accordance with the OMB Information Quality Guidelines²⁷, Primary Data Stewards should take the following steps prior to data dissemination to ensure objectivity and utility of PEPFAR-supported data:

- Prepare a draft of the data set or document after consulting the necessary stakeholders, including government and non-government sources, as appropriate;
- Determine and assure the accuracy and completeness of source data;
- Determine the expected uses by the government and public;
- Determine necessary clearance points;
- Determine where the final clearance decision shall be made;
- Determine whether peer review would be appropriate and, if necessary, coordinate such review;
- Obtain clearances; and
- Overcome delays and, if necessary, present the matter to higher authority.

7.3 Integrity

Several administrative layers ensure the integrity of PEPFAR data against unauthorized, unanticipated, or unintentional modification. Together these controls protect the security of PEPFAR information (see Data Security above).

PEPFAR data systems employ role-based access to prevent unauthorized modification of data. Several key data streams are governed by the PEPFAR Data Calendar (see Data Access above), which promotes integrity by setting clear expectations for when systems are open or closed for data entry and revision.²⁸

7.4 Requests for corrections

When PEPFAR receives information from the public that raises questions about the quality of the information it has disseminated, PEPFAR duly considers corrective action. The purpose of any such corrective action taken is to serve the genuine and valid needs of PEPFAR without disrupting PEPFAR

²⁵ Federal Data Strategy, Principles. <https://strategy.data.gov/principles/>

²⁶ Data.gov. A Primer on Machine Readability for Online Documents and Data. <https://www.data.gov/developers/blog/primer-machine-readability-online-documents-and-data>

²⁷ Executive Office of the President. Office of Management and Budget. Office of Management and Budget Information Quality Guidelines. https://www.whitehouse.gov/sites/whitehouse.gov/files/omb/assets/OMB/infoereg/igq_oct2002.pdf

²⁸ PEPFAR Data Calendar, <https://datim.zendesk.com/hc/en-us>

processes, and to deal with information quality matters. “Affected” persons are those who may benefit or be harmed by the disseminated information. This includes both: (a) persons seeking to address information about themselves or about other persons to which they are related or associated; and (b) persons who use the information.

Persons seeking to correct information affecting them that was publicly disseminated by PEPFAR may submit a Petition for Correction addressed to SGACPublicAffairs@state.gov or:

Office of the U.S. Global AIDS Coordinator and Health Diplomacy
U.S. Department of State
SA-22, Room 10300
Washington, DC 20522-2210

7.5 Data quality roles and responsibilities

7.5.1 PEPFAR Interagency Data Systems Committee

The Interagency Data Systems Committee may provide strategic guidance on data quality.

7.5.2 Program Results and Impact Monitoring for Epidemic Control

PRIME authors PEPFAR-specific data quality guidance and manages the administration of the PEPFAR Data Hub, PEPFAR SharePoint, DATIM, Panorama, and other systems.

7.5.3 Program Quality

Program Quality authors SIMS programmatic content, in collaboration with Agencies, and manages data quality in relevant S/GAC systems and products pertaining to SIMS, in coordination with PRIME.

7.5.4 Agencies

Agencies implement both PEPFAR-specific procedures and their respective information quality guidelines.^{29 30}

7.5.5 Country teams

Country teams are responsible for following pre-dissemination review procedures.

7.5.6 Implementing partners/grantees

PEPFAR-funded recipients provide accurate data in machine readable formats according to the required reporting calendar for each data stream. PEPFAR-funded recipients also perform data quality audits to ensure compliance with data quality guidance.

7.6 Documentation and metadata

Per OMB guidance, PEPFAR implementing agencies maintain an enterprise data inventory that includes data assets created, collected, and managed through PEPFAR.³¹ Data stream-specific documentation

²⁹ USAID. ADS 578, Information Quality Guidelines. <https://www.usaid.gov/sites/default/files/documents/578.pdf>

³⁰ U.S. Department of Health and Human Services (HHS). Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated to the Public. <https://aspe.hhs.gov/report/hhs-guidelines-ensuring-and-maximizing-quality-objectivity-utility-and-integrity-information-disseminated-public>

³¹ Executive Office of the President. OMB Memorandum M-13-13 Open Data Policy-Managing Information as an Asset. <https://www.whitehouse.gov/sites/whitehouse.gov/files/omb/memoranda/2013/m-13-13.pdf>

should include introductory materials that orient consumers to each of the data streams. Authoritative data sets should have accompanying metadata. Metadata should include the following elements: Title, Description, Tags (Keywords), Last Update, Publisher, Contact Point, Unique identifier, Public Access Level, License, and Theme.³²

Known issues are documented in data release notes that accompany data sets. A data dictionary should accompany data sets to support appropriate use of PEPFAR data.

³² U.S. Chief Information Officer. Project Open Data. Metadata Resources for Schema v1.1.
<https://resources.data.gov/resources/dcat-us/>