REPORT

Recommendations on the **Ethical Use of Novel HIV Data & Analytics**

Working Group on the Ethical Use of Novel HIV Data and Analytics

May 2024

Funded by the Bill and Melinda Gates Foundation

ethics lab

GEORGETOWN UNIVERSITY
Center for Innovation in Global Health

COOPER/SMITH

BILL & MELINDA GATES foundation

WORKING GROUP MEMBERS

Janet Tatenda Bhila
Zimbabwe Young Positives (ZY+)

Aleny Couto, MD
Ministry of Health, STI and HIV/AIDS, Mozambique

Shona Dalal, PhD
WHO, Dept. of Global HIV, Hepatitis, and STIs

Will Fleisher, PhD
Georgetown University, Center for Digital Ethics

Jen Gennai
Google, Responsible Innovation

Rayid Ghani
Carnegie Mellon University, Machine Learning Dept. and College of Information Systems and Public Policy

Peter Godfrey-Faussett
London School of Hygiene and Tropical Medicine

Melissa Goldstein, JD
George Washington University, Milken Institute School of Public Health

Kathy Hageman, PhD, MPH
CDC, Division of Global HIV, AIDS, TB

Nina Hasen, PhD
Global Health Consultant, ACHIEVE Innovations

Micheal Ighodaro
Prevention Access Campaign, Washington, D.C.

Thoko Kalua, MBBS, MSc
Center for International Health, Education, and Biosecurity (Ciheb), University of Maryland

Jennifer Miller, PhD
Yale School of Medicine

Yogan Pillay, PhD
Department of Health, South Africa

Anton Pozniak, MD
Chelsea and Westminster Hospital, HIV Medicine; London School of Hygiene and Tropical Medicine

Miriam Rabkin, MD, MPH
Columbia University, ICAP Global Health, Systems Strategies

David Ribes, PhD
University of Washington, Dept. of Human Centered Design & Engineering, Data Ecologies Lab

Lisa Singh, PhD
Georgetown University, Massive Data Institute, Dept. of Computer Science

PROJECT TEAM

Charles Holmes, M.D., MPH, Co-PI
Georgetown University, Center for Innovation in Global Health

Maggie Little, Ph.D., Co-PI
Georgetown University, Ethics Lab

Heidi Weimer, J.D., MPH, LL.M, Project Manager
Georgetown University, O'Neill Institute for National & Global Health Law

Dylan Green, MPH, Subject Matter and Technical Expert
Cooper/Smith, Research and Science

Alicia Patterson, Ph.D.
Georgetown University, Ethics Lab

Jonathan Healey
Georgetown University, Ethics Lab

Sydney Luken
Georgetown University, Ethics Lab

INTRODUCTION

Key priorities of HIV programmes include improvements in the ability to identify those most at risk of acquiring HIV, experiencing treatment interruption, or in need of more support to remain on treatment—all in service of providing access to appropriate services. In support of these goals, there is increased interest by HIV researchers and programme innovators in deploying the new capabilities offered by Big Data, including data generated by individuals' use of digital services and devices, such as social media platforms, apps, mobile phones. In order to leverage the power of Big Data, HIV programmes are using Machine Learning (ML) and Artificial Intelligence (AI) methods to develop models. Advances in Natural Language Processing (NLP) models has also led to interest in extracting information about patients' personal life experiences that may be contained as clinician notes in their Electronic Medical Records (EMRs).

While these strategies have strong potential for improving predictive models, they also bring with them critically important ethical questions. These include issues around privacy and justified surveillance, risks to individuals and groups should the data that is harvested and analytically generated become known, and the potential for algorithmic bias. It is thus critical to appreciate the distinct ethical issues posed by these new modes of surveillance and analysis, and consider frameworks for their responsible use.

International guidelines provide important ethical principles relevant to these questions, including guidance on surveillance in public health ethics,1 the use of AI in health contexts,2 and data privacy and protection.3 The purpose of this guidance is to move from broad principles to pragmatically identifying and confronting considerations, complexities, and factors in the context of HIV. Its goals are to 1) help guide researchers and programme innovators in building designs that are ethically, as well as technically, feasible; 2) assist those charged with assessing proposals for funding and implementation; and 3) identify actions HIV funders and multilateral organizations to advance responsible approaches to the use of novel data and ML models in HIV research and programmatic innovation.

This guidance is the product of an 18-member international, interdisciplinary, and intersectoral Working Group, supported . The Working Group convened remotely for three workshops. Each workshop extended across three days; with half-day sessions. Members included HIV researchers and programme innovators with experience in novel data and machine learning; technical experts from computer, information, and data sciences; experts in global public health policy and health law; experts in data ethics and ethics of AI; and members from the community of people living with HIV who have been leaders in advocating for their community. Countries represented were Malawi, Mozambique, South Africa, the United Kingdom, the United States, and Zimbabwe. A series of briefings, extensive virtual table-top exercises, break-out groups, and anchoring case studies informed extensive and iterated plenary discussions.

This guidance is the product of an 18-member international, interdisciplinary, and intersectoral Working Group. The Working Group convened remotely for three workshops. Each workshop extended across three days, with half-day sessions. Members included HIV researchers and programme innovators with experience in novel data and machine learning; technical experts from computer, information, and data sciences; experts in global public health policy and health law; experts in data ethics and ethics of AI; and members from the community of people living with HIV who have been leaders in advocating for their community. Countries represented were Malawi, Mozambique, South Africa, the United Kingdom, the United States, and Zimbabwe. A series of briefings, extensive virtual table-top exercises, break-out groups, and anchoring case studies informed extensive and iterated plenary discussions.

The guidance proceeds as follows. Part II provides a background on the nature and forms of novel data and accompanying analytics of interest to HIV programmatic needs. Part III presents the five illustrative scenarios that were used to anchor discussion across the workshops. Part IV outlines the ethical considerations, concerns, and challenges the Working Group identified and probed in these efforts. Part V provides a list of seven concrete factors that the Working Group concluded, based on these discussions, were critical to the responsible use of novel data and ML analytics in the context of HIV research and programmatic innovation. Part VI provides concrete recommendations to funders and multilateral HIV organizations for immediately actionable efforts that can make an outsized difference in advancing responsible approaches to this area. Finally, a legal briefing is provided in an Appendix.

TABLE OF CONTENTS

BACKGROUND

The Need for Improved Data & Analytics to Support HIV Programmatic Priorities

Substantial progress has been made toward countering the HIV epidemic. In the past decade, HIV incidence has fallen by one-third and AIDS-related mortality has decreased by half.[4] This progress is due to advances in both HIV treatment and prevention responses, including the improved effectiveness and tolerability of antiretroviral treatment (ART) and expanded access.[5] In addition, combination prevention strategies have supported those at risk of infection: educational programs and outreach, combined with biomedical interventions—such as the provision of condoms, needle exchanges, opiate substitution therapy, voluntary medical male circumcision (VMMC), treatment as prevention (TasP), and oral pre-exposure prophylaxis (PrEP)—have contributed to the decline in HIV incidence rates.[6]

With these advances in hand, UNAIDS set forth diagnosis and treatment goals known as the "95-95-95" targets in 2020 that were adopted by the UN General Assembly in 2021. The goal is for 95% of all people living with HIV (PLHIV) to be diagnosed, 95% of those diagnosed to be on ART, and 95% of those on ART to be virally suppressed, by 2030. In addition, UNAIDS has established aggressive targets for reductions in HIV incidence and increased coverage of prevention programs by 2025.[7]

A critical component in pursuit of these goals is identifying individuals, communities, and areas most in need of treatment and prevention resources.[8] This includes identifying communities and individuals at higher risk of acquiring HIV, so that appropriate prevention resources can be effectively engaged;[9],[10] identifying those living with HIV who are not yet linked to treatment, in order to determine where community resources or new clinics are most needed;[11] and identifying those who are facing challenges staying on treatment, in order to offer them more intensive support.[12]

These efforts include granular questions such as:

- **Which neighborhoods and venues should receive new mobile HIV testing site units?[13]**
- **Which population(s) should be prioritized for combined prevention (including PrEP) and via what venue?[14]**
- **Where should a new ART facility be constructed?[15]**
- **Among ART patients who have not refilled their medication, who should be traced for re-linking to treatment?[16]**
- **Among those who appear to have interrupted ART, who has simply moved to another health facility for care?**

Yet gaining adequate data to inform these programmatic efforts has proven to be a significant challenge. There is increasing recognition that current methods of data collection, such as population-based surveys and aggregate data from program monitoring, have inherent limits for capturing the sheer complexity and highly dynamic nature of the risks at issue.

Complexity of Risk Factors for Acquiring HIV

Much of current efforts have focused on identifying members of demographic groups that have heightened risk of acquiring HIV. These groups include men who have sex with men, people who inject drugs, and, in many countries, adolescent girls and young women, who are often disempowered with respect to sexual encounters.[17] It is increasingly recognized, however, that these demographic factors capture only a portion of

an individual's total risk of infection. Much risk is due to more personal, highly complex, and quickly shifting factors.

For instance, an individual's risk is affected by factors such as the background rate of HIV in their immediate area. But this is complicated by individuals moving from rural to urban areas, experiencing separations from family, and the shifting surrounding prevalence of HIV in the community in which that individual lives at any given time.[18] Short-term, long-term, and seasonal migration, common throughout many African countries, make it difficult to identify the actual risk an individual faces at a given time.

An individual's risk of acquiring HIV is also strongly affected by their own social or sexual network. Yet it has been difficult to ascertain an individual's social and sexual networks. This information is inherently highly personal, private, and hidden, and can change daily. Not only would assessing data on networks be prohibitively expensive, it is unfeasible to obtain the sample sizes and frequency that would be needed to isolate complex risk factors.[19]

Also challenging is identification of local geographical areas with high concentrations of activities that carry higher risk of HIV transmission.[20] National HIV programs, NGOs, and donors attempt to catalog these hotspots so that they can better reach individuals with appropriate peer-led prevention interventions from civil society organizations. Such "hotspots" can be as large as a neighborhood or as granular as a venue, such as a highway rest stop, a brothel, or a dance club.[21] Data to identify these hotspots are currently manually gathered in the field by NGOs. This highly resource-intensive method of collection limits the reach of such efforts, and means the data become quickly outdated.

Complexity of Risk Factors for Becoming Virally Unsuppressed

For those living with HIV and on treatment, risk factors for becoming virally unsuppressed are similarly complex subject to rapid change.[22] Being virally unsuppressed is problematic for both the person's own health and a risk for any exposed sexual or injecting partners. HIV programs aim to identify those who have a heightened probability of being (or becoming) virally unsuppressed in order to direct outreach and bespoke treatment support services to keep people connected to care.

But risk factors for all of these challenges are highly varied and changing, with individuals moving in and out of risk. For instance, risks for becoming non-adherent, or lost to care, include challenges in accessing medication, problems experienced with side effects, changes in relationships, moving to an unfamiliar community, and worries about personal safety should their use of medication be discovered.[23]

Complexities Around Finding, Reconnecting, and De-Duplicating

The concerns summarized thus far can be described as inferential and predictive statistical questions regarding patterns, associations, and correlations. However, there are also more procedural, statistical, data, and programmatic challenges. Many settings in Africa lack comprehensive birth and death registration or universal identifiers. Further, as above, African settings can have exceptional levels of mobility—including short-term commuting, longer-term seasonal migration, and permanent migration.

The absence of vital registrations data and great mobility lead to considerable managerial problems for HIV programs. For example, simply tracking patients on treatment as they move from facility to facility—for various reasons—is a concern for ensuring adherence and viral suppression. Importantly, there are significant differences in the follow-up required for those simply visiting a new health facility and those who have interrupted their ART.

This problem is further exacerbated in the prevention landscape where programmatic systems are still in their infancy and individual level data are nearly non-existent due to technical and stigma-related reasons. It is therefore currently impossible to understand individual-level patterns and trajectories in prevention service uptake, which are critical to continuous program improvement and targeting.

Current sources of information are ill-equipped to gather data that will capture these kinds of complex,

variable, and quickly changing factors. Population-based surveys are highly resource-intensive, and quickly outdated, conducted perhaps once every five years.[24] Program monitoring data efforts provide aggregate information such as the number of people tested, pregnant women treated, or males circumcised[25], and are intermittent—usually collated on a monthly or quarterly basis.

Because of these limitations, there is increasing interest in leveraging the use of Big Data and AI to inform decisions around these key programmatic priorities.

Big Data

The rise of computational power, as well as the enormous expansion of digital devices and services, has led to a revolution in data gathering. "Big Data" is a general term used to refer to the large and complex data, ncluding those made possible by new digital products, that are enabled by the recent massive increase in computing speed and storage.[26] Several features are distinctive of Big Data.

Massive size

As its name indicates, Big Data refers to datasets that contain an enormous amount of information. Sometimes, hundreds of thousands of people or more are represented in a dataset. For instance, Electronic Medical Records (EMRs) are increasingly used for entire patient populations, with clinical data ranging from vitals to laboratory results and imaging. Other times, the number of people might be limited, but the number of characteristics or variables about each person is extensive. For example, because genomic testing is expensive, data is available on a relatively small number of people. However, such testing provides, at a minimum, 3 million variables for each person.

Such size often brings with it a significant increase in the number of predictive variables that can be analyzed.

Novel sources

Some Big Data comes from the digitization of data that had previously been manually gathered, such as EMRs. Big Data, however, is often harvested from expansion in the source of information that is due to new technologies. These include biometric sensors in phones and watches, social media posts and network interactions, and ubiquitous photo and video capture—whether from smart phones and cameras on the ground or satellites in orbit. The rise of social media and the use of smartphone applications alone have led to an explosion of data on individuals' behaviors, beliefs, and networks. These platforms capture data from users, including the content of messages, the location of the user while using the application, and the time that posts or messages are sent.

Such data affords the ability to passively harvest information, rather than relying on the active collection of data, expanding both the types and scope of data that can be collected.

Rich information

Next-generation databases usually contain both "structured" and "unstructured" information. Structured data are relatively simple pieces of information, such as dates, addresses, or a patient's vital signs, that are stored in a predetermined dataset format. Examples of unstructured data, by contrast, include videos posted to social media, digital images such as X-rays, and notes scribed into a digital record. These data are contained in large collections of files that are structured in their native formats. It is estimated that over 80 percent of data that exists today is unstructured.

The richness and variety of these data carry the potential to provide greater insight into, for instance, sensitive sexual and network behavior compared to surveys.

Real-time availability

Big Data has high "velocity." This refers to the speed and frequency of data availability. While manual paper-based systems are slow, sometimes taking months to become available, digital formats of Big Data can be available in near real-time. Traditionally, most data were collected in an infrequent manner—for example, as a part of a survey or a clinical trial. Increasingly, novel data are being captured in a continuous or ongoing manner. Heart rate measurements that are usually collected only annually (if at all) during a routine primary care visit can now be continuously monitored with smart devices.

This velocity brings with it the ability to track rapidly changing factors relevant to HIV, such as mobility, relationship, and hot spot patterns.

AI & Machine Learning

Big Data is, in most instances, unwieldy data. This can limit the feasibility of traditional statistical methods to extract desired information. New methods from computer scientists have developed new systems for analyzing Big Data, and using that data to construct models to be more powerful, perform certain tasks more efficiently, and sometimes more accurately, than can be achieved by traditional statistical methods.

Artificial Intelligence (AI) is a broad umbrella that refers to the ability of computers to emulate human thought and perform tasks in real-world environments. A key ingredient in any AI system is the use ofMachine Learning (ML) algorithms. These are a class of algorithms that enable computers to search for patterns in data by rapidly evaluating thousands or millions of options for which data in the set might be most informative for the relationship sought. In contrast to traditional statistical approaches, in the machine learning context, computers can evaluate thousands of approaches considering virtually the entire parameter space for selection of the optimal approach.

There are a variety of ML algorithm types, appropriate to different types of datasets. All proceed, however, by using the chosen ML method first on "training data"—an initial dataset of interest that has known variables and a value (or label) for the prediction of interest—and then testing the proposed models on new sets of data that were not used to construct the ML model to see how well the model performs on "unseen" data using metrics determined by the data researcher.

As with traditional statistical models, ML-derived models are only as good as the data they are based on. In one famous example, an ML model was trained to classify images of animals and people. The model ended up having a high rate of classifying certain people as wolves. This occurred because all the wolf images the algorithm saw were in the snow, and the model was classifying images with snow as "wolf."[27] It later classified any person—or anything else—on snow as a wolf as well, because of the bias in the training data set. Failure to understand how the ML-model uses information to produce inferences makes it difficult or impossible to assess such models for validity, reliability, or bias.

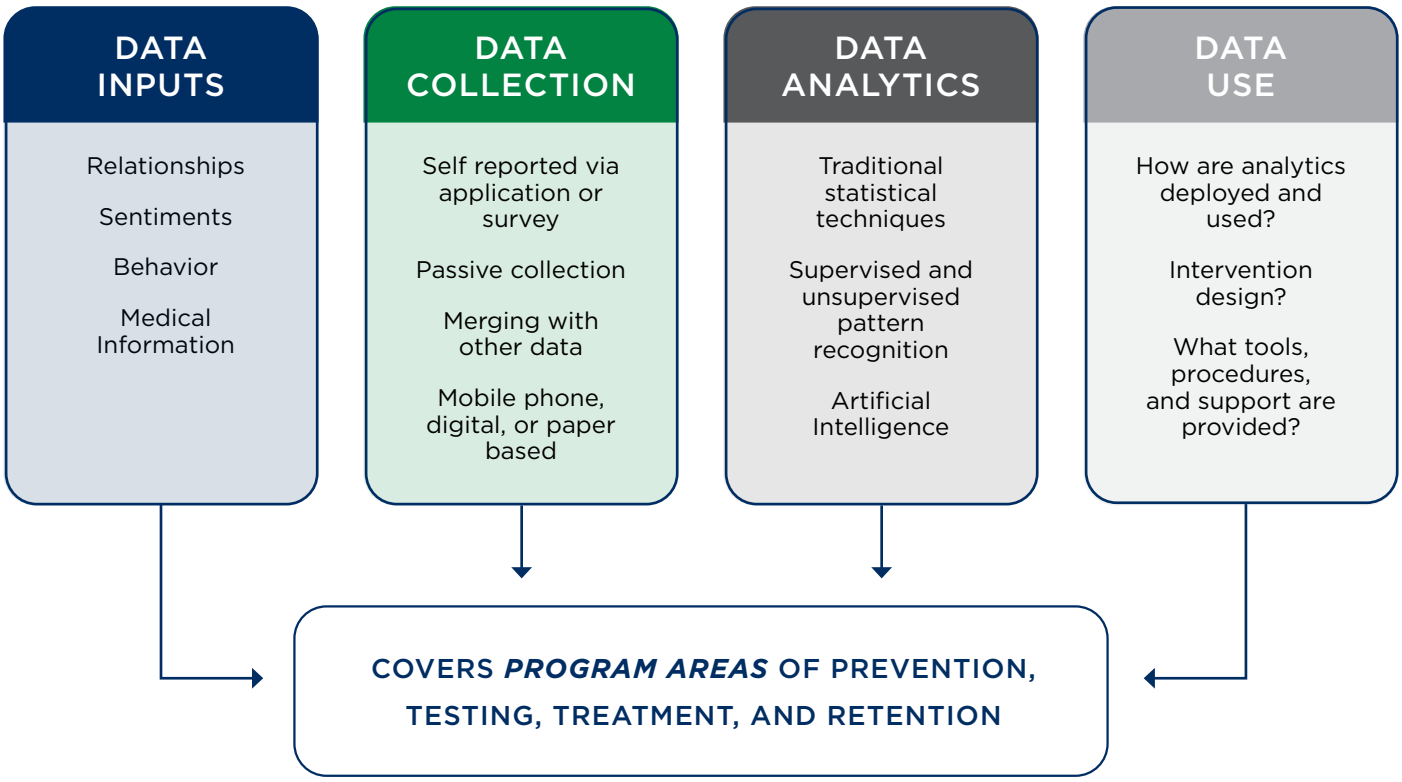
But carefully developed ML-based models can be remarkably successful. In health fields, algorithms have been deployed to evaluate clinical scans from X-rays, MRI, or CT for diagnosing patients with numerous conditions. Models have been successfully used for prediction in hospital settings where the importance of triaging emergency cases or detecting early signs of hospital-acquired infections and sepsis can be lifesaving. Data from EMRs have been analyzed in research projects to deepen understanding of clinical pathways.[28] The use of ML in COVID research and response was prevalent. A living systematic review has identified over 230 published ML prediction models for COVID diagnosis, risk, and prognosis.[29]

Together, Big Data and ML models offer strong advantages. Passively generated, rapidly updated, and rich data-sets expand the size, scope, and granularity of data that can inform programmatic decisions. Once developed, such models are far less resource-intensive than traditional data collection, and reduce inaccurate information that is introduced by, for instance, survey methods that rely on respondents' answers. ML methods, in turn, can extract informative patterns from these huge and unwieldy datasets that are beyond the practical use of traditional statistical methods.

For these reasons, HIV researchers and programmatic innovators are increasingly looking to harvest Big Data, including from novel sources, to develop ML engines to aid in programmatic efforts. Models have been used, for instance, for directing general screening services for HIV testing programs.[30] Models have also been used to develop simple risk scoring algorithms or nomograms to enroll high-risk individuals on PrEP.[31]

ILLUSTRATIVE SCENARIOS

As useful as such approaches may be, they also raise ethical issues that are as complex as they are urgent. To anchor the Working Group's analysis in real world scenarios and contexts, five illustrative scenarios were used. The scenarios are illustrative of the types of projects currently underway or under active discussion. They were chosen to represent examples in each of the HIV programmatic priority domains: prevention, link to treatment, return to care, and retention. These scenarios, supplemented by examples shared by Working Group members, were used to help surface and identify key elements, values, and ethical issues that were then subject to discussion.



X/Twitter Data to Identify Hot Spots

Concept

An important tool for directing targeted HIV prevention efforts is identification of geographical areas with high concentrations of behaviors known to carry heightened risk of HIV acquisition. Current efforts to identify such "hot spots" are labor intensive, which limits the number of areas that can be assessed and means the data become quickly outdated. This scenario involves a government's public health authority finding and using patterns in local social media activity to help identify hot spots across the country in a way that is rapidly updatable.

Research

Researchers would work with Twitter to collect, anonymize, and analyze data from local Twitter accounts set to public. The contents of users' posts would be analyzed using Natural Language Processing (NLP) methods. These are models that allow a computer to extract the meaning of key phrases in a text, including categorizing text that fits "features of interest" supplied by the programmer—for instance, behaviors known to carry heightened risk of HIV acquisition. The output of the NLP stage would be combined with the location and time stamps of each post, the frequency of tweets coming from given accounts, as well as external data on currently known hot spots. Together, these data would serve as the training data for a machine learning model to look for "archetypes" of tweets that are associated with sexual behavior carrying heightened risk of HIV acquisition. Programmatically this would inform further hotspot identification and prevention outreach and support.

Program

If a validated predictive model were to be identified, the public health agency or authority (PHA) would launch a program. It would put the handles of all local Twitter public accounts on a secure, private server, and collect the above data on an ongoing basis. It would conduct the above analyses on anonymized posts, and attach labels to tweets that have high predictive probability of behaviors known to carry a heightened risk of HIV acquisition. The analysis would then gather the location data of all tweets carrying that label, and that location data would be aggregated to identify areas with high concentrations of tweets indicating the heightened risk. These areas would be shared as an updatable list with the health authority to inform their programmatic response.

Dating App Meta-Data to Identify Individuals at Heightened Risk of HIV Acquisition

Concept

Dating apps have grown rapidly in popularity in recent years and have been of much interest to the public health community as a potential channel for health communications and promotion of sexual health interventions. These apps are used in outreach programs, such as local HIV prevention services that provide testing, circumcision, and/or PrEP. While helpful, broadcast messaging is also limited in its effect. More effective are targeted messages to those at heightened risk of HIV acquisition that can provide information and access to priority prevention resources.

This scenario concerns the possibility of an NGO working to find and use patterns in user account meta-data to identify users more likely to have heightened risk of HIV acquisition, and then sending targeted messages to these users and to users in high engagement with them.

Research

In this scenario, researchers would send in-app messages to users of a dating app, asking if they would like to take a voluntary, anonymized questionnaire to help assess their potential risk of acquiring HIV. The survey would include questions about the user's age, sex, education, and other demographics, along with questions about certain sexual behaviors, such as their number of recent sexual partners and whether or not they use condoms or PrEP. The questionnaire results would be analyzed using existing, validated risk scores to categorize the individuals as high-, medium-, or low-risk for HIV acquisition based on their responses. Users would have HIV acquisition risk levels attributed to them; the results would be relayed to respondents.

Researchers would also ask these users' permission to harvest meta-data from respondents' anonymized accounts. These include the number of direct messages sent and the number of accounts engaged with, which can be indications of likely in-person meetings and potential exposures; frequency and duration of sessions on the app, which can be indications of how active the user is in seeking partners; and time stamps, since posts sent and received in the very early hours of the morning may reflect activities with a higher risk of transmission. Researchers would also analyze the number of direct messages sent to or received from specific accounts, and would gather the same meta-data on those accounts. A machine learning model would then be trained on these data to develop a model to predict users with heightened risk of HIV acquisition.

Program

If a validated predictive model were to be identified, the NGO would enter an agreement with the dating app company, with coordination with and oversight by the local public health authority, to launch an ongoing program. The NGO would provide the dating app company the protocol for processing data and applying the predictive model, and then apply to the above, anonymized raw meta-data from all in-country accounts on the app. The accounts flagged as having heightened risk of HIV acquisition would receive the appropriate message types supplied by the health program in the app automatically.

EMR Clinical Notes to Prevent Treatment Interruption

Concept

An important priority in HIV programs is to predict, find, and help people who have stopped refilling their prescribed antiretroviral therapies. A wide variety of personal and life circumstances can contribute to treatment interruption, including difficulty tolerating side effects, fear of stigma if their medication were discovered, difficulty getting to a pharmacy, depression, pill fatigue, or a new relationship. Tracking or tracing programs work to identify those with treatment interruption and to connect them to healthcare staff, community organizations, and trained peers who offer support for their re-engagement and return to ART use. While these tracing programs are important, they are both costly and slow, which leaves many with treatment interruption experiencing poorer health and at increased risk of transmitting HIV.

This scenario concerns the possibility of PHA working to develop and then use a model to predict those at heightened risk of treatment interruption, so that support can be offered to help prevent the interruption of treatment in the first place. In particular, this scenario considers the possibility of using NLP models to extract information from clinician notes scribed into patients' Electronic Medical Records (EMRs), which may capture information about personal and life circumstances. The aim of this would be to develop a predictive model that can flag the records of patients who are at risk of treatment interruption so that clinical staff can identify patients who need enhanced support services.

Research

Researchers would ask for permission from health authorities to access the data from a large sample of anonymized EMRs of patients diagnosed with HIV. An NLP model would be used on the clinician notes to extract categories of interest identified by the researcher (examples might include “new relationship” or “feeling depressed”). The NLP-processed text for each patient’s EMR would be added to the clinical data it contains (i.e., height, weight, lab results, history of infections, etc.), including, importantly, any data about subsequent treatment interruption or being virally unsuppressed. An ML model would be trained on this data to see whether there are “archetypes” of EMRs that are predictive of treatment interruption.

Program

If a validated predictive model were identified, then health authorities would launch an ongoing program of processing the clinician notes for all patients diagnosed with HIV, linking the information to their clinical data, and applying the ML model to the combined data. When the predictive model identifies a patient as being at high risk of treatment interruption, that status would be added to the patient’s EMR in order to alert facility staff to offer enhanced support or specific services to the patient to bolster the patient’s continued use of ARTs.

Meta-Data from Push Notification Programs to Identify Individuals at Risk of Non-Adherence

Concept

Push notification programs allow patients living with HIV to opt-in to receiving text reminders to take their medications, refill them at the pharmacy, or remind them for dates of medical appointments. These programs, which are often run by NGOs, allow patients to customize message reminders to keep the sensitivity of the message private in case others see the text.

This scenario concerns the possibility of an NGO developing and deploying a model that uses meta-data of those who are enrolled in their “push notification programs” in order to deliver in-app messages offering care, and also the opportunity to enroll in “assisted partner notifications services” to partners who may have been exposed.

Research

Researchers would ask patients enrolled in push notification programs for permission to access the meta-data on read receipts and response receipts on their anonymized account, and to allow researchers to access the anonymized EMRs to collect clinical data on factors related to adherence. These data would serve as the training data for a machine learning model to predict heightened risk of non-adherence.

Program

If a validated predictive model were to be identified, the NGO would launch an ongoing program to harvest the above meta-data from all of its push-notification accounts, which would remain anonymous. The NGO would apply the ML model and accounts matching the archetype for enhanced risk of non-adherence would be sent messages via the notification program offering resources to support adherence. They would also receive messages asking whether they would be willing to voluntarily enroll in an assisted partner

notification service. Such programs ask those who are virally unsuppressed if they are willing to provide the names and contact information of their recent sexual partners, in order to reach out to those partners to notify them of their possible exposure (without mentioning any names), encourage testing, and provide prevention services or treatment as needed.

Mobile Phone Meta-Data to Identify Hot Spots through Social Networks

Concept

An individual's social or sexual network strongly affects that individual's risk of acquiring HIV. Currently, it is very difficult and expensive to identify and assess individuals' sexual networks, especially at scale and on a continuing basis. This scenario concerns the possibility of a PHA working to find and then use patterns in mobile phone meta-data indicative of an individual's own network having a heightened risk of HIV acquisition. The texts and calls made into and out of those phones would be matched to their location data. Areas with heightened concentrations of calls and texts from and to those phones would be identified as hot spots to which prevention resources could then be directed.

Research

Researchers would ask people at chosen venues, including HIV testing clinics, if they would be willing to participate in a study in which they would be asked to share their (understanding of their) HIV status; take a sexual risk behavior survey; and agree to have their mobile phone meta-data gathered for a period of time. This meta-data would include the number, timing, and location stamps of calls and texts coming from or received by their phone (more specifically, the phone's SIM card); the phones (SIM cards) those calls/texts are sent to/received from; and the corresponding meta-data from those phones' use. All data would be anonymized.

Machine learning methods would be applied to the research data set to see if there are complex patterns in account meta-data that correlate with the reported HIV status and risk scores from the sexual risk behavior survey, in order to develop an "archetype" of an account's meta-data that is predictive of heightened risk of acquiring HIV.

Program

If a validated predictive model were to be found, the government would enter into an agreement with the telecommunications firm or mobile network operator (MNO) operating in-country to perform two steps. The MNO would be provided with the protocol for processing anonymized raw data and applying the predictive model. The MNO would apply the model to the anonymized raw meta-data from all in-country SIM cards, code those at high-risk, and generate and link HIV risk acquisition scores to SIM cards.

Second, the MNO would then harvest the location data of all texts and calls issued or received from high-risk-score SIM cards. The locations of those texts and calls would be aggregated to identify areas with high concentrations of high-risk-score SIM cards. Areas identified as "hot spots" of such activity would be shared as an updatable list with the health authority to direct targeted prevention resources to those areas.

ETHICAL CONCERNS & RISKS WITH THE USE OF BIG DATA & AI IN HIV RESEARCH & PROGRAMMATIC INNOVATION

While the collection and harvesting of large data sets from non-traditional sources and the accompanying ML methods for their analysis hold great promise, proposals for, proposals for expansion into this space brings critically important, and often novel, ethical issues. The Working Group had robust and extended discussion of these issues, based both on the scenarios formally developed and their experiences of similar programs underway, to probe and identify key ethical issues and their contextual complexities.

Expansion of Surveillance into Private Spheres

One of the concerns extensively discussed by the Working Group was privacy. Efforts such as these represent an expansion into new territories of data surveillance.

Privacy & Digital Spaces

The Working Group discussions strongly underscored the core importance of privacy as a value and right. A key aspect of privacy is the importance and value of having control, with respect to certain kinds of information, over who that information is shared with. Retention of "informational boundaries," that is, retaining a level of control over what is known about one by society, specific people, or institutions, is an intrinsic value, important to protecting one's personal data.[32] Also important is having spaces that are free from monitoring, especially ongoing monitoring, without one's consent. Protected spaces enable a broad latitude of actions and self-expression without fear of others' watchful eyes or one's every move being tracked.[33]

Relationships and social interactions increasingly happen in digital, not just physical, spaces. Extending the ongoing collection of personal data from activities on browsers, platforms, and apps can impinge on what have become important sites of personal and social life.

The Working Group also emphasized that privacy is not just an issue about individuals. Just as people wish to have personal spaces free from surveillance, groups can also have an interest in maintaining communal spaces that are safe from monitoring.[34] Further, in the world of Big Data, information gathered on a small number of subjects can affect what is known about many.[35]

While not all digital online users constitute a community in the important sense of the term, some apps, services, and platforms provide places where people do gather as a community. Still others, such as dating apps, are spaces where personal exchanges among members is the aim of the site. A program of harvesting certain personal data that runs on all in-country accounts on the app may be experienced as an important intrusion, changing the very nature of the space.

Factors relevant to the level of concern include specifics on how private users regard the particular digital activity, and how sensitive are the data collected. For instance, the Twitter scenario involves gathering the content of Twitter posts from accounts set to "public view." This brings more modest concern for impingement of autonomy-based privacy interests (though it is not without some level of consideration, as people often have little understanding of the various uses to which such data might be put, and might not approve of certain uses[36]).

At the other end of the scale is the mobile phone scenario. While only meta-data is captured, rather than message content, the meta-data in question is location data, gathered on all mobile phone users on an ongoing basis. Here, users have expressed no individual choice, and the stakes are higher as well. Given the degree to which mobile phone use is integrated into life and the frequency of accessing it, tracking a phone's location data is, for all but sporadic users, a near-equivalent to tracking the individuals.

Chilling Effects

Surveillance into digital spaces can also have chilling effects. It can cause people to avoid participating in activities that are important to maintaining the social fabric. Various apps and platforms, for instance, can function as places for democratic activities such as political activism; overly intrusive surveillance of these platforms can weaken or prevent these activities.

Surveillance may also cause people to reduce their use of programs that are meant to support them. Information included in EMR clinician notes, for instance, can contain stigmatized information, such as reports of domestic violence and of suicidal ideation. Harvesting this data raises privacy concerns, and may have chilling effects on what patients share with their health providers. Similarly, an important question raised for the push notification proposal is whether adoption of this program would decrease participation in these important programs.

Gathering, Generating, & Linking Sensitive Information

Another central issue for the Working Group—sometimes known as "informational risk"— concerned the harms that individuals or groups may face should the information collected from these digital spaces and technologies become known. Importantly, the informational risks discussed focused both on the data that is harvested, and the attributes that are generated in their analysis.

Harvesting Sensitive Data

The first and most obvious concern is that information harvested from novel digital sources can include highly sensitive information, such as, for example, HIV status or risk-related behavior. Information gathered from individuals' use of digital services and devices—from which web pages were accessed, to patterns of cell phone movement—can also include highly personal or potentially compromising information, quite apart from sensitive HIV-related information. Clinician notes in EMRs can contain information such as reports of domestic violence and of suicidal ideation, both of which are highly stigmatized.

Generating & Linking Sensitive Attributes to Accounts

Sources of informational risk are not limited to the sensitive information that is harvested. Informational risk can be introduced when sensitive attributes are generated by the analysis of the data and then linked to individual records or accounts.

The EMR and X/Twitter scenarios provide one helpful example. The projects in these scenarios aim to analyze ordinary language—clinicians' scribing of narrative notes from an office visit, and the content of social media posts, respectively—using Natural Language Processing (NLP). NLP methods include extracting or engineering key "features of interest" supplied by the researcher or found through ML methods. In the case of these scenarios, features of interest might include things such as "engages in IV drug use," "MSM behavior," or "transactional sex." These features represent highly stigmatized and often criminalized behavior, and here are linked to the individuals' records. Such information would be extremely time consuming to retrieve without the NLP-generated attributes; here, it is extracted, interpreted, compressed, and represented in a tag.

More generally, for all of the scenarios considered, the ultimate analytic output of interest are individual risk scores for acquiring HIV, being virally unsuppressed, or non-adherent. Once again, in these scenarios, these generated risk scores are linked to the individual accounts monitored. In the programmatic implementation of the mobile phone scenario, for instance, the record of every mobile phone (more precisely, every SIM card) would have an attribute concerning highly sensitive, and often stigmatized, HIV factors attached to it.

Increased Risk of De-Anonymization

Data collected for public health surveillance purposes are virtually always formally anonymized at key stages. Such anonymization, which has become increasingly sophisticated, involves removing, masking, or encrypting names and other personally identifying features such as addresses, and then assigning a unique identifier to the records. A key concern with big data and their associated analytics, however, is the potential they carry for allowing individuals' identities to be recovered.

Cross-Referencing across Datasets

A person's identity can often be inferred when multiple datasets are merged together, as is common practice in big data analytics—by piecing or layering information together.[37,38] The advent of Big Data has substantially increased the ability to de-anonymize data by this method.[39]

Data can also become de-anonymized when cross-referenced or "triangulated" with other, less anonymized, datasets available from other available sources.[40] An enormous amount of data about individuals exists in a variety of databases, from posts on social media accounts that are registered under real names to public census records. While the specific steps for de-anonymization by cross reference are highly technical and specific to the databases, once those steps have been identified, they are easily shared with others who can execute them at will. In one study, it was found that, in some cases, only 4 data points from one dataset can unlock the identity of an individual within an anonymized one.[41] Furthermore, prolific social media platforms and web footprints allow for record linkage across even anonymized data sources.[42]

As the Working Group emphasized, there is no such thing as true anonymization in the era of Big Data.

Highly Identifying Geolocation Data

De-anonymization can also occur when the data collected is highly identifying. For instance, even a perfectly anonymized dataset of mobile phone call records holds the potential to infer the identity of the phone's user. This is because the data includes precise time and location patterns of the phone—which means precise information about where the phone user was and when. Because people's patterns of movement across time are unique, that information can soon translate into the ability to identify the person behind the movement. For instance, examining patterns in the time and location stamps contained in a data set may reveal instances in which the evening phone activity from a specific SIM card tends to come from the same location almost every day—which is a likely indication that the location in question is the person's home address.

Geospatial tracking of citizens on an open-ended basis therefore raises serious concerns when pursued outside of certain extenuating circumstances, such as designated public health emergencies. In response to COVID-19, for instance, many programs pursued widespread continuous tracking of location data from mobile phones (pursuits which engendered their own ethical discussions[43, 44]). Whatever the resolution of those debates, these programs were instituted in the context of a global public health emergency. That context—and its official designation— triggers different standards of tolerable privacy infringements, with their own strict criteria, including explicit criteria for sunsetting the temporary measures. Because the HIV epidemic is open-ended, this means that the programmatic data harvesting should be as well. This is a factor that must be considered in determining its justification, including consideration of precedents it may set.

Fewer Surrounding Protections

One of the most critical questions around proposals for harvesting and analyzing the data envisioned in the scenarios is who will have, or might gain access to, the data in question. Questions of access and data protection, important for any data collection and retention effort, are particularly urgent when the data in question are from the new digital sources at issue in this guidance. Several distinctive features were underscored in Working Group discussions.

Commercial Partners

Access and analysis of novel data also often involves commercial partners, such as telecommunications, app, or platform companies. Commercial entities have interests that can be at odds with public health interests, including interests in monetizing shared data or outputs of analytic systems shared with it,[45] [46] or in combining the analytic output with their own, often extensive, data on users.

Data protection of such arrangements relies heavily on specifically built agreements and the de facto enforcement power of the government. The latter can depend, in term, on the particular power balances of the companies, such as the telecommunications company and in-country regulators -- a power balance that in some countries may tilt strongly to the former.

Lower Regulatory Protections

More generally, personal data sourced from the use of digital platforms, apps, and devices, whether or not it is sourced from a commercial entity, is subject to far weaker regulatory protection than is health data, especially health data generated in traditional health systems settings.

While the data collected in the EMR and push-notification program scenarios are squarely under regulatory protections for health data, for instance, the data collected in the rest of the scenarios has far less regulatory protections than health data. Data privacy regulations around sharing of and access to data from social media sites is still weak in most countries. While the number of such regulations is increasing (see Appendix on Legal Briefing), their stringency -- and enforcement -- vary significantly across countries and regions, and are in any event far less than regulatory protections of health data. In practice, in many countries the central determinant of access to users' data is simply the company's policy.

Access to mobile phone meta-data, in turn, falls under the regulatory body that governs telecommunications; but those regulations do not afford the same protections as do regulations governing health data.

In short, the move from health data gathered in traditional health settings to these novel sources of data is a move from high to low regulatory protections. This is both why the novel sources of data are so attractive to projects, and why more guidelines for their use are sorely needed.

Potential for Civil Rights Violations

One of the central areas of discussion concerned the potential risks of government access to the datasets and analytic models contemplated. In certain political contexts, governments may have strong motivations to de-anonymize databases and use the information to target individuals or persecute vulnerable populations.

These issues can be especially important to consider in contexts where HIV is stigmatized and certain behaviors relevant to HIV risk are criminalized. Working Group members, for instance, shared experiences in which data collected by NGOs to help identify sex workers as a priority population for HIV prevention outreach and support was sought by police departments seeking to arrest them.

Geolocation data, once again, drew particular scrutiny in discussions. Serious concerns were expressed about proposals that would involve governments' housing or having access to individual-level mobile phone geolocation data. Governments can use the information from ongoing geospatial tracking to find clusters of people fitting a specific profile and determine their concentration. Individuals may not know they are being tracked or profiled in this way, and they may have little recourse with their government.[47]

Even without de-anonymization, aggregates of geolocation data can and have been used in ways that violate civil rights. For instance, Working Group members shared experience of a project in Malawi that found that it was possible to use anonymized CDR data to understand potential mass gatherings and thus COVID super-spreading events there. There was a highly contentious election happening that year. Since mass gatherings could include political gatherings, the data could have been used by the government for political purposes. Here, the risks far outweighed the benefits of collecting and developing inferences from the data, and the project was not pursued.

If the political context is volatile and democratic structures protecting sensitive data are fragile, risks of government access and misuse can be prohibitive. While some governments may be respectful of the ethical and privacy concerns over collecting and analyzing sensitive data, other governmental entities may be less so. Further, incoming regimes may be less considerate. Without safeguards or policies in place to protect data in such political transitions, there are risks of significant civil rights violations.

Finally, even in stable political contexts, there nevertheless remains the important potential for function creep.[48] Once datasets and predictive tools exist, it can be tempting to repurpose them, using them for more extensive aims than the parties involved had originally agreed upon. This can include temptations to redeploy the data and tools for uses, both public and private, that have less oversight, transparency, or legitimacy. These expanded uses may not be vetted or given any specific justification or due process.

Community Involvement

One of the most important issues raised, and returned to throughout the workshops, is the critical importance of community involvement in these efforts.

The Working Group strongly underscored the principle that community involvement is ethically essential for responsible, legitimate, and effective public health surveillance programs. The reasons are multiple and well known. Programs that do not speak to the actual needs or priorities of the community divert scarce resources from their needs, which can be urgent. Programs can bring with them unanticipated harms and costs borne by those they are meant to serve. Initiating programs to help others without giving direct voice to those others can be an expression of arrogance and a form of epistemic injustice.[49] The issue is compounded when the absence of voice is influenced by the very historical patterns that lead to the presence of need. Most central is the principle that those most affected by decisions should have a say in them.

While always important, community involvement is particularly urgent when considering the sorts of programmatic innovations envisioned here, as the degree of intrusiveness and privacy infringement of these emerging possibilities is poorly understood, and cannot be assumed or determined simply by legal or ethical theory. The more distant from settled precedent and previously socially negotiated reasonings for surveillance, and the more uncertainty there is about scope or probability of harmful implications, the stronger the right of the community to have their voices and views of the relative benefits and costs of a proposal at the center of design and decision heard.

Involvement vs. Engagement

As Working Group members from the community of PLHIV emphasized, community involvement is different from community "engagement" or "education." The latter terms can suggest a one-directional relationship of building communities' capacity and understanding, when communities are experts on their needs and values: their involvement adds to technical and programmatic actors' capacities. They can also suggest an

instrumental view of connection to the community to the needs of the program, rather than centering those most affected by decisions as an equal partner in them.

The Working Group emphasized that community partnership and empowerment should extend to all stages in the life-cycle of a program, not just consultation and input at the end of a project in yes/no adoption decisions. Proper community involvement also includes partnership at early stages in the conceptualization and design of a program; in evaluations of updated information about post-launch efficacy and risks; and in decisions about potential changes, mitigations, or decisions to end a program. The community needs to be at the center of the decision-making: if people affected by a policy or project are not a real part of it, the program's justification is in question.

For these reasons, the Working Group urged that projects identify and establish community partners as early in the concept development stage as possible.

Challenges for Community Involvement

Those with experiences of community involvement also emphasized challenges to achieving effective and respectful community involvement.

There was much discussion of the significant challenge of achieving community involvement in light of the highly technical nature of these research and programmatic proposals. The Working Group voiced/believes that there is an urgent need for research and pilots for modes of empowering communities to participate in decision-making around the technical issues involved in these programs.

The Working Group also discussed challenges of determining *which* communities are the relevant stakeholders, and again how to get effective and fair *representation* of those communities. In HIV specifically, stakeholders include not only those who are living with HIV, but those who are at higher risk of acquiring HIV, and groups who are affected independent of individual HIV risk, such as those living in neighborhoods identified as "hot spots," or members of groups identified in the aggregate as having higher predictions of HIV risk. Members of the Working Group emphasized that the process of community involvement may need to be iterative, to increase the chances that the relevant people are meaningfully engaged.

The Working Group also cautioned against modes of community representation that over-index on professionals in that community, who may be less likely to experience directly the impact of programs. Further, requirements for, and funding to support, community involvement, if not thoughtfully designed, can also end up directing resources disproportionately to large NGOs. Large NGOs are more likely to have structures and resources in place -- in part because of prior fundings supplied -- for the practical aspects of community involvement. This structure can end up disadvantaging smaller community organizations, who are often closer to the ground, from receiving awards.

Promising Models

The Working Group shared creative suggestions for what meaningful community involvement, empowerment, and partnership could look like. One member surfaced the idea of giving this challenge to the community, without antecedently identified proposals for solution, and ask what they would do to address the problem. In thinking about designs for effective community partnership, that is, the relevant community itself could be charged with designing how they will themselves be engaged in and help to govern a particular project.

Ideas surfaced included building in community-partnered reassessment, after community co-design, of programs, to ensure that unintended harms experienced by the community, and again benefits, have input into changes needed. Grants to community organizations to develop their own proposals, design feedback mechanisms, and community governance are important measures that could expand support for community-led design.

Access, Data Protection, & Data Governance

As emphasized in the WG, a core ethical principle of data ethics is: "Do not collect if you cannot protect." There will be limits to the levels of protection that can reasonably be expected: trade-offs will always exist between protection and gaining data that can be ethically important in its own right. But data collection, it was emphasized, brings with it strong obligations of protection - of due diligence, technical structures, policies, and enforcement.[50] The higher the informational risks, the higher the standard of required efforts. If meeting the standard proves infeasible, this is a limiting condition on the ethical justification for a data collection project.

Assessment of data protection and governance includes 1) assessment of the entity envisioned to house and access the harvested and analytically derived data, and 2) what technical and policy structures would be in place to guard against expanded access or misuse.

Structures of Planned Access

Structures of planned access often involve a mix of commercial partners, NGOs, and government agencies. Roles include who will own and access the *raw data*; who will perform the *preliminary feature building* -- and hence have access to its generated attributes; who will apply the *predictive model* to the processed data -- and hence have access to the individual-level risk score information; and what form of data is ultimately shared with *implementation partners* -- the individual-level data or aggregate-level data constructed from them.

Different arrangements can bring with them different kinds of factors and levels of concern. For instance, the push-notification scenario has the advantage of involving a non-commercial app that has been commissioned by the NGO. All elements of the program -- the data collection, application of the ML model, and analytic output of interest (the code for risk of non-adherence), and delivery of the messages -- stays within the non-commercial app.

The dating app scenario, in contrast, involves an arrangement between an NGO and a commercial entity -- the dating app company. In this scenario, the NGO provides the predictive model and content of the messages; the commercial company applies the model to their raw data the app is generating, and retains access to the analytic output.

The scenario about mobile phone meta-data, in turn, involves a cooperation agreement between the MNO and government. In this scenario, the MNO processes the raw meta-data they collect in-house, then attaches the analytic output to SIM cards' anonymized IDs, and shares aggregated summaries of HIV risk within a geographical area to the public health authority, rather than the individual-level data from which they are derived. This structure has the strong advantage of sequestering the potentially compromising geolocation and other data from mobile phone use from the government. That said, it may raise questions about how the MNO might end up using the data and analytic output. The success of the sequestering, in turn, can depend on details of the relation between government and MNO.

Data Protection, Governance, & Monitoring

Data governance refers to the policies and technical structures that govern access, use, and security of databases to protect the data against misuse. Data governance is implemented via data governance boards, who dictate rules for access and use.

The Working Group discussed real-world pressure such boards may experience. It emphasized the importance of making use of institutional and technical structures that may help it protect the data from potential incursions. Examples discussed included the use of data enclaves, which sequester data on private servers, and the use of data intermediaries -- neutral institutional entities that can serve as gatekeepers

to data access. The Working Group also emphasized the need, when assessing what data protection and governance structures and oversight should be instituted, to engage in scenario analysis, identifying and contingency planning for potential outcomes that may arise, including changes in the political context or ruling party.

Data protection for ongoing programs also includes establishment of structures for their monitoring and oversight. Lifecycle approaches emphasize that responsibilities for data with programs do not end upon their launch; in many ways, this is where those responsibilities begin.

Algorithmic Bias & Complexity in ML Models

In addition to ethical issues raised by expanded data collection from new digital spaces and technologies, there are important, distinctive ethical issues attendant to Machine Learning models.[51]

Algorithmic Fairness

One area of deep importance to discussions of Machine Learning algorithms is fairness. Data used for training and validation can reflect ethnic, socioeconomic, differential access, or other structural disparities.[52,53,54] [55] Using that data to guide future decisions may reinforce those disparities. For instance, a model developed for Covid-19 aimed to predict geographic areas where vaccine "acceptance" was greatest, to prioritize the deployment of scarce vaccines to areas where they would most likely be used and not expire or be wasted. The risk is that what was actually predicted was not "acceptance," but rather greater access to care, education levels, and socio-economic status. Using the model to allocate more vaccines to these areas (where acceptance appeared greatest) furthers disparities and enriches this already advantaged population.[56,57]

Similarly, adding clinician notes from electronic medical records may, when combined with other data, give insight into risks of treatment interruption. But the extensiveness of clinician notes might be biased in favor of those in higher resourced clinics, where nurses have more time for patient discussion and note taking. Generalizations on this data may end up directing resources disproportionately to those already privileged in the health system, worsening inequities.

Interpretability

Machine learning models can be extremely complicated. With some ML methods, computer scientists can probe the model to discover which variables turned out to be predictive. Many ML techniques, however, including some of the most powerful, are sufficiently complex that even those who designed them do not fully know how their predictions are made. These models are sometimes referred to as "black box" models.

Lack of explainability matters ethically. Because it is difficult to determine how each prediction is made, uninterpretable systems can make it much more difficult to spot socially biased results. Lack of interpretability can also impede the ability of communities to appeal consequential decisions made on their basis. If individuals or communities believe a decision regarding their treatment was unfair or wrong-headed, and those to whom they would appeal cannot interrogate the basis of the decision even with the help of computer science specialists, it can make challenging the decision or seeking redress all but impossible.[58]

Even models that are interpretable by experts may not be easily understood by the people affected by them or by those making adoption decisions. This can limit the ability of non-specialists to understand the system they are considering adopting, and of community advocates from assessing and having a voice in those decisions.[59]

Diversifying Optimization Metrics for Ethical Values

The above issues about bias and interpretability underscore the fact that technical choices made in the development of ML models can have ethical consequences.

The Working Group urged the importance of diversifying optimization metrics (the specific measures of success that programmers select to instruct how the ML algorithms will choose among options for determining archetypes). It was shared in the Working Group that computer scientists have historically focused solely on metrics of accuracy, power, and efficiency. While these features are clearly important, transparency, fairness, and interpretability are also key values. Some of the most powerfully predictive models can also carry the most ethical costs. It may be appropriate to sacrifice some degree of accuracy or statistical power in favor of approaches that minimize or reduce ethical risk.

The Importance of Auditing ML Models

There was extensive discussion of the importance of auditing ML models that are being proposed for deployment -- especially large-scale, consequential deployment, as with the HIV programmatic innovations considered in this guidance. Audit of ML models includes assessment of the datasets used as training data, the choice of analytic methodology, and the optimization metrics used. Access to this information is hence critical to assessing claims about the models' accuracy and generalizability.

For this reason, it was noted that use of models published in academic research papers, or those sold by commercial software companies, should be approached with caution. Often, access to the data sets used to train the model, or the optimization metrics used, is not available. This means that the model cannot be assessed for whether there were options to add fairness- and privacy-preserving metrics.

Research Design, Review, & Informed Consent

Assess for Social Value

The Working Group cautioned against the pursuit of "research for research's sake" when considering research with big data and analytics in the HIV context. Given the substantial costs of such research, and the informational risks it may carry, such research should only be undertaken as a testing ground for a programmatic innovation concept, actively in consideration for implementation, that would, if successful, carry strong social value.

Critically, the ethical nature of a proposed program is relevant to judgments of its social value. assessment of a program's potential social value must include assessment of its ethical nature. If the proposed program - the ongoing harvesting of data it envisions, the actor(s) that would house or have access to the data, etc. -- is ethically unjustified, then the social value of the research, and the significant investment of resources it will require, is likely unjustified as well.

ERC/IRB Review of Research Proposals

Research conducted on data harvested from sources such as social media platforms and apps is sometimes assumed to be exempt from ERC/IRB review, as research that involves no appreciable risk to the data subjects. However, theWorking Group emphasized that the kinds of research at issue with the scenarios should require ERC/IRB approval, as these protocols involve linking sensitive data to individual records.

The social media protocol is a helpful example. While the social media data is only harvested from accounts set to 'public,' the protocol involves generating risk scores of acquiring HIV or being virally unsuppressed, and linking those attributes to accounts. In the dating app scenario, even if harvesting data from the

app were otherwise viewed to be exempt from ERC/IRB review, the protocol links that data to surveys -- traditionally gathered research that is subject to such review.

There was a strong consensus that the research protocols envisioned for all of the illustrative scenarios should be subject to formal IRB/ERC review and require informed consent.

Informed Consent

As with research in other highly technical areas, it is challenging—though no less important—to meet the relevant standard of understanding and appreciation of what one is actually agreeing to and its attendant implications and risks. It is challenging to describe not just the nature of the data collected but the attributes that the analysis may ascribe to the subject (such as an individual score for risk of HIV acquisition), the relevant possibilities of other agents who might have—or later gain—access to the data, and the implications such access could carry. As with any research in highly technical areas, design and testing of effective communication, along with community outreach, will be essential.

The Working Group was deeply cognizant of challenges in achieving the adequate understanding needed for formal consent. Significant questions about whether the bar for informed consent can be met: whether people can really understand what they would be agreeing to, and its implications for their lives, the inferences that might be drawn about them, and risks of de-anonymization. Working Group members underscored the challenge in reaching this criteria when highly technical issues are involved, as with novel data and analytics. No solutions were forwarded; the Working Group regards this as another critical area for investment in study and development of potential models of effective explanations and engagements with potential participants.

The Working Group also cautioned against asking for consent for permanent ongoing harvesting. A model of recurrent versions of consent and data usages, checking back In on use of data, as well as control to opt back out of a study, can take into account life circumstances change, or the person changing their mind.

Challenges with ERC/IRB Capacity

The Working Group also identified the important deficit in IRB/ERCs experience and needed expertise to review these specialized research protocols. It is difficult to submit protocols to IRB/ERCs if they lack available expertise to assess them. This challenge is a common one across countries, as the rapidly escalating sophistication of these analytic methods has far surpassed the more usual clinical or social science expertise currently populating IRB/ERCs. The limit on this capacity is especially significant, and urgent, in many sub-Saharan African countries. This highlighted the urgent need for increasing training of IRB/ERCs, and increasing the availability of expert consultants that can help in the review of these protocols.

Ethical Assessment of Programmatic Proposals

While it is generally well understood that health research requires ethical review, the concept of formal ethical evaluation is less established for programmatic innovations. However, the Working Group discussed the critical need for ethical evaluations of proposals for programmatic innovations of the sort considered in this guidance.

The ethical stakes of harvesting personal data in the research context are fundamentally different from those of ongoing, passive harvesting for a programmatic effort. The former is limited in scope and duration, proceeds (or should proceed—see Part V, Section V) only with the informed consent of data subjects, and involves testing, rather than use, of the predictive model developed. The programmatic effort, in contrast, involves population-wide harvesting of personal data, usually without consent or opt-out conditions, in an ongoing and usually open-ended manner, with real-world decisions being made on the basis of the predictive analytics.

It is critical that funders of programmatic innovation projects, and government adopters, should have processes in place to review and evaluate these proposals. Also critical is for such processes and structures to include community representatives.

To enable such evaluation, it is important that proposals provide clear descriptions of the features that have high ethical salience. Concept briefs for research and programmatic proposals around data and analytics can tend to focus on elements relevant to technical and institutional feasibility, sometimes in great detail. But they may not always include, explicitly address, or highlight factors that are needed to assess the proposal's ethical feasibility. As discussion of the cases deepened with the ethical deliberations, the Working Group articulated the importance for proposals to clearly state:

What specific *data* will be collected;

From what *sources*;

Under what *disclosure and choice conditions* for data subjects;

***Owned and accessed* by whom;**

With what additional analytically *generated attributes* linked to individuals;

For what *analytic output*;

Using what *analytic* method;

***Shared* with what programmatic implementers and in what form;**

Overseen by what *data governance structure*;

Developed with what *community involvemen*t and plans for future involvement;

With what provisions for *lifecycle* monitoring and capacity to respond.

TOWARD ETHICALLY RESPONSIBLE USE OF BIG DATA & AI IN HIV RESEARCH & PROGRAMMATIC INNOVATION

Based on the discussion, preceding, identified key priority. In navigating the ethical values at stake in the use of novel data and ML methods, research and programmatic innovators, as well as those who serve as funding or adoption gatekeepers, should address the following processes, standards, and expectations.

Embed Ethics Into the Concept Development Process

Concept development is a time of option scanning, early scoping, and preliminary feasibility assessments. While the technical, methodological, and institutional feasibility of options are critical, so too is their ethical feasibility.

Options raised should be probed for potential ethical risks from the start. Questions include how intrusive are the contemplated means of monitoring; how consequential would it be if the data collected or analysis generated were accessed; how readily might the data be de-anonymized, and who might have interest in doing so; whether the concept carries substantial risks of public misunderstanding or worsening of trust in public health efforts; whether the country's political context raise concerns for the possibility that the dataset or ML model might be misused against its citizens.

The intersection of technical, methodological, infrastructural, and ethical considerations is both complex and specialized. Project teams are strongly encouraged to convene a joint advisory group that combines technical experts in data science and ML methods, an ethicist fluent in the technical and public health issues, along with community members briefed on the technical and ethical stakes. Early identification of such a team can help at the concept development phase and beyond, as projects inevitably face recalibrations, refinements, and iterations.

Incorporate Robust Community Involvement

Community involvement in consideration of public health programmes - always critical - assumes particular importance programmes proposed involve expansion into new and less tested areas of digital surveillance.

Those who would be most affected by initiatives should have opportunities, not simply to comment on whether or not to accept a given proposal, but to influence the shape and direction of a proposal, to give feedback to ideas and offer their own—to ensure that priorities and goals, and ideas for how to meet them, are shaped with and by the community's perspectives.

Proper community involvement thus includes partnership at early stages in the conceptualization and design of a programme; in evaluations of updated information about post-launch efficacy and risks; and in decisions about potential changes, mitigations, or the decision to end a programme.

Project teams are encouraged to consult and adapt existing models that have successfully developed community-led approaches. These include building in community-partnered reassessment after community co-design to ensure the scope and changes will meet the community requirements; and giving grants to community organizations to develop their own proposals, design feedback mechanisms, and community governance.

Design Value-Centered ML

Programmatic innovations should aim to use interpretable methods wherever possible. If and where such methods are unviable, researchers should consider sacrificing some accuracy for the sake of increasing interpretability. Proposals for using uninterpretable models should be assessed, not just by the accuracy advantage they may bring, but whether that advantage is worth the added risks of bias and the social opacity they bring.

Those developing ML models should measure bias using fairness measures and reduce any model bias that may arise. Active areas of computer science research are focusing on developing privacy- and fairness-preserving techniques that should be explored. The addition of fairness in optimization metrics should be considered.

Projects should not use ML models that they cannot audit—both for accuracy and the values internal to the system. Audits should include understanding the training dataset used, optimization metrics deployed, and be assessed for alternatives that may represent a better overall balance of values. Audits should also be conducted to make sure performance is consistent over time after models are deployed in real world settings, and that they do not diminish over time to the point of limited utility (in which case ethical risks obviously outweigh benefits).

Ensure Robust Data Protection & Governance Structures

Given the sensitivity of the data involved and its linkage to individual accounts, data protection and governance structures must be especially robust.

Data governance boards should be designed in ways that protect the independence of the board and its ability to resist potential pressures for access. This includes attention to appointment policies, accountability audits, and checks and balances in the governing structure. It may also involve sequestering data on private servers, and/or use of neutral data intermediaries that can serve as gatekeepers to the data.

Systems must have robust structures for monitoring, oversight, and capacities to respond with mitigation should unintended consequences or risks emerge.[60] Plans should include regular assessment and re-evaluation points, with processes, responsibilities, and authorities demarcated for response. Policies should think ahead to conditions under which the data structure should be destroyed, and develop mechanisms and protocols for doing so before crises emerge.

Systems should retain data and analytical engines only as long as they are required for immediate programmatic needs. To drive down the ethical risk around the long-term life of these data sets, researchers and programmatic innovators should consider amassing data sets that would be designed to be destroyed after their intended use. Programs should also assess whether it is essential that an analytical engine persists indefinitely, or whether it should be a time-limited tool that self-destructs or is decommissioned by a certain date.

Require ERC/IRB Review for Research

Research that harvests personal data from extant sources should be subject to IRB/ERC approval when it will be linked with other, formally gathered research data, or when it will be linked to other sensitive data, including sensitive data that is analytically generated.

Informed consent for research participation requires the provision of clear explanations of what the research involves, its burdens and risks; and its purpose. When describing the study and potential risks, care must be taken to go beyond description of the immediate data that will be harvested, and include inferences that

can and will be drawn from it, risk scores that will be linked to their individual records, and the risk of de-anonymization.

Assessment of the research proposal should look ahead to the ethical feasibility of the programmatic proposal to which it is in service. If the proposed program—the ongoing harvesting of data it envisions, the the actor(s) that would house or have access to the data, etc.—is ethically unjustified, then the social value of the research, and the significant investment of resources it will require, is likely unjustified in turn.

Require Ethical Evaluation of Proposals for Programmatic Innovation

Proposals for programmatic innovations should be subject to ethical evaluation before adoption. Considerations to be addressed include the topics outlined in Part III of this guidance: autonomy burdens and informational risks to both individuals and groups; implications for trust and adoption of needed services; elements of bias, transparency, and accuracy for the ML model.

Attention must also be paid to the political context in which the research and programmatic innovation would take place. Tools justified in one context can be used for oppression and violations of civil rights in another. While those who build analytic tools cannot always foresee or control who may make use of their tools, concrete risks of oppressive uses and civil rights violations must always raise questions about the responsibility of pursuing the tool's development.

Evaluation of proposals should also address the comparative advantage of the proposed approach for achieving expected health outcomes relative to other, potentially less ethically risky or intrusive options. Considerations should be given to the anticipated data protection, governance, and monitoring plans. Particular focus should be given to the choice and design of ownership and access structure; details of any share-back arrangements or vulnerabilities they possess should be disclosed. Consideration should also be given to what community involvement was incorporated into the design and development of the programmatic proposal.

Evaluation of the proposed innovation should involve advisors or consultants knowledgeable about the technical details, programmatic experts, as well as community representatives—all of whom should be treated as equal partners.

Ensure Public Disclosure & Accountability

A core tenet of public health ethics is the need for public transparency, justification, and accountability of public health surveillance programmes.[61] [62] [63] This obligation applies both to governmental and large non-governmental entities working for public health goals. The public has a right to know what personal data is being monitored, collection, and analysis of personal data. accountability for effects, harms, and responsible reviews; and to give input into discussions of benefits and harms.

Public health authorities and non-governmental organizations engaged in the use of novel data collection and analytic methods must give meaningful public disclosure of ongoing data harvesting, including what data is being collected, the use to which it will be put, the reasoning and policies that lie behind its decisions, the results of monitoring and mediation activities; and provide meaningful opportunity for public input and objection, with special attention to those most directly impacted by the programmes.

Channels for transparently sharing program details with the public will vary between public health authorities and NGOs. The level of detail provided will vary according to the scale and import of a program. At a minimum, both PHAs and NGOs should disclose ongoing data harvesting and the purpose for which it is gathered. Adding disclosure to the terms and conditions of a surveilled digital platform or app does not

qualify as public health disclosure. Program sponsors should be open and welcome to inquiry and scrutiny, including from peer organizations and, especially, community advocates.

RECOMMENDATIONS FOR HIV MULTILATERAL ORGANIZATIONS & FUNDERS

The following recommendations outline concrete and immediately actionable recommendations for HIV funders and multilateral organizations to help advance responsible approaches to the use of novel data and ML models in HIV research and programmatic innovation.

Support Community Agency

There is an urgent need for research and pilots of models of supporting meaningful community involvement in consideration of these potential projects. Multilateral HIV organizations and funding partners should support the development of programmes and translational tools, co-designed with community representatives, that would help to empower communities to participate in decision-making around these programmes.

Funders should support development of translational tools that would help to empower communities to participate in decision-making around proposals for novel data analytics programs. Tools should be developed by working groups that include not only technical content experts, but those with expertise in community-centered design, and members of the community.

Funders should consider funding the training of community representatives and advocates to help support the creation of a generation of voices equipped to defend the interests of their communities vis-à-vis the increasing use of data collection and analysis. Funders should also consider leveraging the methods that researchers for HIV vaccines and cures use to inform their workshops and symposia that actively put community advocates in the room with researchers.

Funders should consult with relevant community organizations and advocacy groups to explore what support would be most helpful to promote their ability to influence and assess proposals for the use of novel data and ML models to guide resource distribution decisions. Questions to ask these groups might include what translational tools might be most helpful.

One such example might be a “Smart Questions” tool, which would be co-designed with community advocates, to help translate technical aspects of proposals and their stakes for the community. Such a tool provides key questions for community members to ask of proposals to empower communities to participate in decision-making around the technical issues involved in these programs. Such a tool would also help such groups to confer with and advocate with respect to donors and government agencies. Tools developed for community inclusion in research, such as those outlined in Good Participatory Practice (GPP) Guidelines, may be helpful in building further models for community inclusion in programmatic design and decisions.

Increase Expertise in Settings and Organizations where Solutions Are Deployed

Multilateral HIV organizations and funding partners should support briefings and capacity building for Ministry of Health staff on the technical and ethical issues surrounding data collection and analysis efforts, in order to enhance governments’ ability to independently assess proposals.

Funders should consider developing programmes to help in-country researchers and programmatic

innovators to become a resource for their countries by increasing their ability to initiate appropriate projects, and their ability to assess projects proposed by others. Such work might include piloting summer schools that provide education on the technical and ethical issues surrounding the use of Big Data and ML models for up-and-coming in-country researchers.

The rapidly escalating sophistication of these data harvesting and analytic methods has surpassed the more usual clinical or social science expertise currently populating IRB/ERCs. Funders should consider developing training modules for in-country ERCs on the technical and ethical issues surrounding these systems, to increase ERCs' capacity to provide effective oversight of research proposals, including when outside experts should be consulted in the review process.

Establish Processes for the Ethical Evaluation of Programmatic Funding Proposals

Funders of programmatic projects should develop and institute processes and procedures for structured ethical evaluation of proposed projects.

In support of this aim, funders should require provision of a formal ethics risk assessment at the full proposal stage. Such assessments should address identification of privacy concerns, including risks of informational harms, de-anonymization, and civil rights violations; and identification of potential bias or problematic uninterpretability of the ML model. They should include identification of political or power-based concerns that increase the potential for misuse of the data or analytic engine, and whether adequate community involvement was incorporated into the design and development of the programmatic proposal.

Proposals should provide clear, accessible descriptions of features relevant to ethical assessment, in language accessible to reviewers from different disciplines of public health, HIV, data science, ethics, and computer science, along with community representatives involved in the review process.

Project proposers should also be required, as a condition of funding, to incorporate ethics/safety planning and implementation activities throughout their projects. Funders should prioritize projects that have a clear ethical risk assessment and decision-making processes outlined throughout their timelines.

Funders should require, and provide resources to enable, effective community inclusion in projects that they fund. Funders should consider requiring that a community expert be involved in the project as a technical advisor from the project's beginning. Such representatives—as with any other experts—ought to be financially compensated for the time they are asked to dedicate to ongoing projects. Dedicated funding for community involvement should be built into the grant mechanism.

Finally, funders should consider giving preferential funding to proposals that incorporate robust and innovative community partnerships across the lifecycle of their projects, including community organizations close to the ground, and strive to make funding available to enable those partnerships.

Expand Ethical & Technical Support for Researchers & Programmatic Innovators

The novel data harvesting and analytic methods at issue in this guidance present a combination of technical and ethical risks and issues that many teams are not immediately equipped to handle. Because of this, it is important to support the inclusion of ethics and technical advisors/consultants or team members that project teams may identify for their own proposals.

That said, for many teams, it can be difficult to identify and recruit experts on their own. Multilateral organizations and their funding partners should work to develop access routes to such experts. Examples could include establishing a centralized group of advisors and consultants with the relevant expertise to

serve as a resource for proposal developers. Experts in these domains with experience in developing these sorts of complex projects (and not just those who specialize in end assessments)are especially valuable.

Multilateral HIV organizations and funding partners should also support development of templates, models, and tools to assist project teams in implementing the suggestions in this guidance. Examples could include a model of, or a suggested process for, developing a robust data governance plan, and an updated risk assessment tool that is commensurate with the wide reach of AI systems. Such an assessment tool would help diversify the types of risks considered, including by assessing risks at both community and individual levels, risks by time scale (e.g., short, medium, and long term), and risks which are "one off" versus cumulative.

Convene International Meetings

Major HIV policy institutions and funding partners, such as the World Health Organization, Joint United Nations Programme on HIV and AIDS, The Global Fund to Fight AIDS, Tuberculosis, and Malaria; and the United States President's Emergency Plan For AIDS Relief (PEPFAR), should convene meetings relevant to their remits to bring attention to the ethical issues that surround the use of novel data and ML models. Meetings of key representatives from the community, implementers, government, and the technology sector can further develop suggestions, recommendations, and normative guidance for reducing ethical risks.

For instance, the WHO's Department of Global HIV, Hepatitis and Sexually Transmitted Infections Programmes could hold a scoping meeting on the subject of Big Data to put the topic on their radar—and on the radar of all who look to the Department for guidance. UNAIDS could, for instance, gather its country and/or regional teams to provide case study examples of where Big Data is already used in the fight against AIDS and where its resources would be most welcome. UNF's DIAL could use its expertise to present ethical exemplars from the projects they support.[64]

APPENDIX: LEGAL BRIEFING

This briefing provides an overview of notable legal rules and laws governing data protection. Included are summaries of rules and laws from global entities, such as the United Nations; Europe; the African Union on Cybersecurity and Personal Data; Kenya; Nigeria; Rwanda; Uganda; South Africa; and the United States.

Heidi R. Weimer, JD, MPH, LLM

The digital era of the last twenty years has ushered in a slew of new laws and rules pertaining to data protection. The Covid-19 pandemic recently has prioritized data protection as an urgent issue for policymakers worldwide as they have been forced to grapple with legalities and ethics of data collection and processing for public health purposes. Thus, within the last couple years, even countries and regions with previously few to no laws pertaining to data protection have since begun to draft and enact data protection measures.

Today, 137 (71%) of 194 countries have implemented domestic legislation protecting data privacy to at least some extent.[65] Roughly half were enacted in the last 10 years; one-fourth in the last 5 to 7 years.[66] By 2023, roughly two-thirds of the world's population will have personal data protection via privacy regulations.[67] Some regulations are sectoral in nature (i.e., specific to an industry or field) while others are omnibus (one or more national laws with multiple sectoral regulations within).

While most of the regulations share basic underlying principles (e.g., respect for individual privacy), specifics vary from one jurisdiction to the next. This makes it difficult for organizations to comply with all applicable rules (especially when rules reach across borders), but also highlights the importance of an ethics guidance framework that can be applied globally. In fact, as laws are rapidly changing, an ethics framework can be vital in influencing and informing policymakers as to what specific actions pertaining to data protection are most ethical and most practical, and thus most sensible. At the same time, legal frameworks can provide the enforcement that ethics frameworks lack.

GLOBAL INSTRUMENTS

No single, global agreement devoted to data protection binds all countries. However, the United Nations (UN) considers privacy protection a fundamental human right—vital to protecting individual freedoms, freedom of expression, and personal dignity. This fundamental right is explicit in the following international legal instruments.

Universal Declaration of Human Rights

The Universal Declaration of Human Rights (UDHR)[68], adopted by the UN General Assembly in 1948, establishes specific rights and freedoms for all human beings. These rights include the right to privacy. Article 12 of the UDHR states that "[n]o one shall be subject to arbitrary interference with his privacy, family, home or correspondence, nor to attacks upon his [honor] and reputation. Everyone has the right to the protection of the law against such interference and attacks."

UN International Covenant on Civil and Political Rights

The International Covenant on Civil and Political Rights (ICCPR)[69], a multilateral treaty which went into effect in 1976, commits member states to uphold specific rights of individuals, including the right to privacy. Article

17 of the ICCPR defines the right to privacy in identical terms as the UDHR (see above). The right to privacy enshrined in the UDHR and ICCPR covers all aspects of an individual's life, including their personal data, as such data might reveal information concerning family life, religious beliefs, political affiliations, and sexual preferences.

International Health Regulations

The International Health Regulations (IHR) is a 2005 treaty legally binding 196 countries and creating rights and obligations of those countries concerning public health events.[70] While the IHR primarily concerns reporting and responses of states and the World Health Organization (WHO), the agreement also addresses safeguards to protect personal data, informed consent, etc., when applying the health measures outlined in the IHR. Article 45-Treatment of Personal Data requires that data be processed fairly, lawfully, and not excessively; be stored accurately and only as long as necessary; and be kept up-to-date, with the ability to erase or rectify.[71]

UNGA Resolutions – The Right to Privacy in the Digital Age

Within the last decade, the UN General Assembly (UNGA) issued two resolutions concerning the right to privacy in the digital age.[72] These resolutions—which carry weight under international law—urge member states to review their privacy laws and practices and make modifications as necessary in order to uphold individuals' right to privacy.

UN Report: Principles Underpinning Privacy and the Protection of Personal Data

This July 2022 report[73] by the UN Special Rapporteur on the right to privacy promotes the protection of personal data and privacy as a fundamental human right in that it enables the protection of other fundamental rights. As such, individuals have the right of protection over sensitive data (e.g., health information), disclosure of which could jeopardize a person's fundamental rights. Sensitive data, then, should be processed only if lawful and necessary, and where the principles of consent, transparency, fairness, proportionality, minimization, quality, responsibility, and security can be upheld.

EUROPE

With the most stringent data protection laws in the world, the continent of Europe is the paragon for data protection laws worldwide, serving as a model for other jurisdictions.

Convention 108 / 108+

The Council of Europe (CoE), composed of 46 member countries, is the leading international organization for the protection of human rights on the European continent. The CoE upholds the right to privacy as a fundamental human right.

The CoE's Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data (Convention 108) (first signed in 1981) is the first legally binding international treaty focused on data protection and privacy.[74] The treaty is global (open to all countries), voluntary, and mutual (i.e., not unilateral); it provides for reciprocal obligations as opposed to an imposed standard placed on countries. Convention 108 requires state parties to update or enact domestic legislation to uphold the protection of fundamental rights with regard to processing of personal data for individuals within that state's territory.

Convention 108+ (2018)[75] is the modern version of the original treaty and today is the only binding legal agreement on data protection. The treaty currently has 55 state parties. Convention 108+ applies in both the

public and private sectors and includes guidelines on big data and artificial intelligence as well as specific measures for health data. Relying on mutual commitment and cooperation, the treaty allows countries to work together to uphold the right to privacy of individuals while guaranteeing data protection to the highest degree possible.

General Data Protection Regulation

The paragon of data protection regulatory frameworks worldwide, the General Data Protection Regulation (GDPR) (effective May 2018) is a comprehensive data privacy regulation which is intended to strengthen individual rights. The GDPR is broad in nature and harmonizes data protection laws across Europe.[76] Reaching beyond the borders of Europe, the regulations govern how entities obtain and process personal data of Europeans, applying not only to organizations located within Europe, but also to entities outside the territory if data subjects are EU citizens. Under the GDPR, states are required to implement or update domestic legislation that either aligns with or surpasses provisions in the GDPR. Non-compliance results in hefty fines for violators.

Basic rights of data subjects under the GDPR include the following: the right to be informed, right of access, right of rectification, right of erasure (right to be forgotten), right to data portability, right to object, and rights related to automated decisions. In addition, entities must produce a privacy policy and designate a data protection officer (DPO). If processing large amounts of data or sensitive personal data, organizations must implement data protection impact assessments (DPIA).

The GDPR replaced the previous Data Protection Act of 1998 (DPA) but is undergirded by the same 6 privacy principles as the DPA.[77] These principles can assist data processors and controllers in determining whether their data practices are permitted by the GDPR.[78]

- **Lawfulness, fairness, and transparency**: Data processors cannot process data simply because the data exists. Instead, they must identify a lawful basis for processing data[79]: *consent* (has been obtained and is clear and affirmative); *contract* (e.g., agreeing to terms and conditions and providing contact information); *legal obligation* (e.g., to assist law enforcement); *vital interests* (i.e., subject's life is dependent on it); *public task* (e.g., public duty or exercising official authority); or *legitimate interests* (e.g., fraud prevention). In addition, data must be processed in "good faith" (i.e., belief that it is honest and legal) and with transparency (informing data subjects how their data will be collected and processed, outlined in a required privacy policy).
- **Limitations on purposes of collection, processing, and storage**: Data must be "[c]ollected for specified, explicit and legitimate purposes" and cannot be stored indefinitely.
- **Data minimization**: Entities may collect only and no more than the data needed. Data that is no longer needed must be deleted.
- **Accuracy of data**: Data should be accurate and up-to-date because accuracy ensures privacy (e.g, If a wrong phone number is contacted, confidential information could be disclosed to the incorrect person.). Data subjects have the right to rectify or erase data.
- **Data storage limits**: Data cannot be stored indefinitely nor in a form that allows reidentification for any longer than necessary. If data is no longer needed for processing, it must be anonymized or erased immediately.
- **Integrity and confidentiality**: Data must be protected against unauthorized or unlawful processing, damages, breaches, etc., by whatever means necessary and technologically available (e.g., encryption, pseudonymization/anonymization, etc.).

In addition to the principles above, the GDPR relies heavily on accountability,[80] with the data controller being ultimately responsible for GDPR compliance.

Recent EU Developments

- **Introduction of data trusts**: A data trust is a new mechanism for holding data where data is pooled and rights concentrated in the trust, such that the "power inherent in aggregated data is returned to the individual data subject."[81] Data trusts are particularly relevant to research and open source data, allowing data to be accessible as needed while still protecting privacy and the rights of data subjects.

- **EU Data Governance Act**:[82] This recently proposed (2020) framework enables the using and reusing of data for the public good. The Act promotes "data altruism" and aims to promote the sharing of beneficial data by "ensur[ing] an effective means of allowing public, 'supervised' access to data sets that have scientific research value."[83]

- **EU Schrems cases**: Two recent EU Court of Justice (EUCJ) cases highlight the importance of data protection and reinforce the notion that EU levels of protection in the GDPR travel with the data. In *Schrems I* (2015), the EUCJ held that the transfer of data of European data subjects (Facebook users) was no longer permitted via the EU-US safe harbor framework. As a result, the new EU-US Privacy Shield was established, imposing more stringent obligations on US organizations processing EU data. However, in a subsequent case, *Schrems II* (2020), the EUCJ invalidated the Privacy Shield and instead upheld the use of Standard Contractual Clauses to guarantee safeguards and uphold rights of data subjects. Further, the decision obligated contracting parties and data protection authorities to assess the adequacy of protection in a recipient country. To be considered "adequate" as a recipient of EU data, appropriate safeguards, enforcement, and legal remedies must be available either in domestic law or international agreements, even when standard contractual clauses exist between parties. In addition, the entire legal regime of a country will be considered when determining adequacy of that country as a recipient of EU data.

- **EUCJ and Grindr**: In 2022, the EUCJ ruled that entities must protect "adjacent data"—i.e., data that might indirectly relate to or could reveal sensitive information (such as health, sexual orientation, etc.).[84] In the original case, Norway had fined Grindr $6.7 million after public officials in Lithuania had their sensitive data disclosed when their spouses' names were published online, because the names could indicate the sexual orientation of the official. The EUCJ's ruling essentially expanded the definition of "sensitive data" to *potentially* sensitive data; thus, otherwise non-sensitive data that can be used to *infer* sensitive data about an individual should itself be treated as sensitive.

AFRICA

The decisions of the EUCJ in the *Schrems* cases (requiring a determination of "adequacy" in a recipient country), the ability of the GDPR to reach beyond Europe, the rapid rise of digitization, and the Covid-19 pandemic have recently encouraged jurisdictions worldwide to enact legislation that aligns with GDPR standards and principles. Nowhere is this phenomenon more apparent than in Africa, where 61% of countries now have some type of data protection laws.[85] Unlike the GDPR's emphasis on data subject rights, however, most African countries' laws focus on empowering national governments to define data privacy for themselves.

Convention of the African Union on Cybersecurity and Personal Data

This 2014 African Union (AU) Convention (also known as the Malabo Convention) establishes a normative framework, broadly covering data protection across Africa in order to harmonize data and privacy protection laws and standards.[86] The framework aligns with social, economic, legal, and cultural environments across the continent and aims to balance the use of technology, handling of personal data, and the protection of privacy with the free flow of data transnationally.[87]

Specific provisions of the Convention include the following:

- Article 14(6)(a): Data is not to be transferred to non-AU countries unless the receiving country can ensure adequate protection over the data (though "adequate" is not defined).
- Article 9(1)(c): The Convention applies only to the "processing of data undertaken within the territory" of an AU member state.

The AU, however, though instrumental in promoting data protection standards, lacks the enforcement power of the EU, as the AU is an intergovernmental organization as opposed to a supranational political organization like the EU. Thus, the Malabo Convention is merely a model framework setting "aspirational" standards for member countries, encouraging them to enact domestic law and regulations that align with those standards. Further, the Convention has been ratified by only a few of the 55 AU countries.[88] Until the Convention sees more widespread ratification, an ethics framework can prove useful in helping to set specific standards for organizations handling personal data.

Nonetheless, even if they have not ratified the Convention, many African countries are enacting data protection legislation. For example, South Africa, Kenya, and Nigeria have not signed the Convention but have implemented robust national privacy laws. Several other countries are currently drafting data protection legislation which likely will be enacted in the next few years. These countries include Eswatini, Ethiopia, Malawi, and Zimbabwe, among others.[89] For many African countries, these new data security and protection regulations reflect their own Constitutional guarantees of the right to privacy as a fundamental right.[90] Additionally, many recently drafted national laws across Africa mimic the GDPR in underlying principles and requirements.[91]

Kenya

Kenya's Personal Data Protection Act of 2019 (DPA)—the country's primary data protection law—serves as a comprehensive legal framework for data protection. The law gives effect to Articles 31(c) and (d) of the 2010 Kenyan Constitution, enshrining the right to privacy of every person, including "the right not to have information relating to their family or private affairs unnecessarily required or revealed or the privacy of their communications infringed."

The DPA establishes guidelines on the use of personal data and creates the Office of Privacy Commissioner (ODPC) for enforcement of the law. Similar (though not identical) to the GDPR, the DPA establishes the following rights of data subjects: the right to be informed of the use of their personal data, access to the data, object to processing in whole or in part, correction of data, and the deletion of false or misleading information. Data controllers and processors are required to obtain consent from data subjects, limit the retention of data, register with the ODPC, conduct impact assessments if there exist potential risks to the rights of data subjects, and report any breaches. Penalties may include fines and imprisonment. In addition, data subjects may seek redress and compensation for any damages and distress caused by violations of their rights regarding personal data.

Nigeria

Reflecting a national trend to increasingly protect data subject rights and strictly enforce the laws in place (including the right to privacy in Nigeria's Constitution), Nigeria's robust Data Protection Regulation (NDPR) of 2019 is the country's primary data protection law. Originally, the law placed enforcement and regulatory power in the National Information Technology Development Agency (NITDA). In February 2022, however, the federal government established the Nigerian Data Protection Bureau (NDPB), which is now the regulatory authority for data protection in Nigeria.

Modeled after the GDPR in its rationale and scope, the NDPR outlines the rights of data subjects over their personal data, including the right to information in easily understandable writing, to withdraw consent at any time, to be informed of safeguards in place when information is transferred across borders, rectification of inaccurate data, data portability, and confirmation of processing, among other rights outlined in NDPR

Articles 2 and 3. Data controllers and processors must ensure consent, designate a data processing officer, ensure proportionate processing of data for purposes intended, risk and impact consideration, and the ensuring of minimal risk of interference with fundamental rights. Further, data processors must submit audits to NITDA. Penalties include fines and imprisonment.

In January 2023, the Nigerian Federal Executive Council (FEC) approved the Nigeria Data Protection Bill of 2020, which will establish a unified regulatory framework for personal data protection. The new law, which is more comprehensive than the NDPR, awaits passage by the National Assembly.

Rwanda

Rwanda's 2021 Law on the Protection of Personal Data and Privacy (Data Protection Law) gives effect to the fundamental right to privacy established in Article 23 of the Rwandan Constitution, which dictates that the "private life, family, home or correspondence of a person shall not be subjected to arbitrary interference; his or her honor and good reputation shall be respected." The Data Protection Law mimics both the GDPR and the Malabo Convention in the underlying principles of lawfulness, fairness, transparency, purpose limitation, and accuracy, and obligations such as the authority for regulators to hold non-compliant organizations accountable (including the imposition of fines). The law places regulatory authority in the National Cyber Security Authority (NCSA) and businesses and organizations to designate an individual data protection officer, provide impact assessments, and notify of data breaches. The law applies to residents of Rwanda (even if not citizens) and any data processing that takes place in the country, as well as to those outside of the country who process the personal data of subjects within Rwanda. Processing and storage of data is allowed only when the data processor holds a registration certificate authorizing such storage. Penalties include administrative fines.

Uganda

Uganda's Data Protection and Privacy Act of 2019 and the Data Protection and Privacy Regulations of 2021 give effect to the right to privacy in Article 27(2) of the Ugandan Constitution, which establishes that "[n]o person shall be subjected to interference with the privacy of that person's home, correspondence, communication or other property." The laws regulate the access, collection, processing, and transfer of data and creates the National Information Technology Authority (NITAU) for enforcement. The NITAU maintains a national register of all public entities and bodies that collect or process personal data. Notably, Uganda's data protection law aligns with the Universal Declaration of Human Rights.

South Africa

South Africa's Personal Information Protection Act (POPIA) of 2020 applies to all entities that collect, store, process, and transmit or transfer personal data for business purposes.[92] The Act creates an independent Information Regulatory Authority that is responsible for investigating, monitoring, and enforcing compliance with the POPIA throughout the country. The POPIA prohibits the processing of "special" personal data (such as race, health, sex life, biometrics, etc.) unless for one of the following authorized purposes: consent from the data subject, research purposes, or care and treatment of the subject.[93] Like the GDPR, the POPIA contains an adequacy requirement, such that data protection travels with the data.

Though modeled after the GDPR, the POPIA differs in some respects that could potentially limit the public good of the data.[94] For example, many bioethicists have concerns about the POPIA because, although the GDPR makes certain exceptions for research and allows for secondary use in specific cases, the POPIA makes no such exceptions. In other words, the South African law does not allow for data subjects to provide broad consent of use over their data, and instead requires that personal data be collected only for a "specific, explicitly defined and lawful" purpose and that subjects must be aware of said purpose. Bioethicists point out that it is counterproductive to prohibit data subjects from providing broad consent over their own data, when that very data could ultimately benefit the vulnerable communities from which many of them come. The conundrum created by South Africa's law is a perfect example highlighting the importance of bioethicists speaking into the drafting of legislation regarding privacy data.

UNITED STATES

The US has no single overarching data protection law. Instead, the US utilizes a "patchwork" approach with sector-specific regulations and state laws. While the EU has a privacy-first agenda, the US is more "hands-off," an approach that tends to favor companies and corporations handling data.

Fair Information Practice Principles

The Fair Information Practice Principles (FIPP) are a set of principles US federal agencies use in evaluating individual privacy in systems, processes, and programs.[95] Though not mandatory, these principles have been helpful in informing the drafting of federal and state laws in the US and have influenced the laws of other countries and the practices of some organizations. The FIPP include access and amendment, accountability, authority, minimization, quality and integrity, individual participation, purpose specification and use limitation, security, and transparency.

Federal Trade Commission Act

The Federal Trade Commission Act (FTCA) allows for data use except when such use is unfair or deceptive.[96] The FTCA covers health information collected by a commercial company (such as Fitbit, Apple, Ancestry.com, etc.) but does not apply to non-profit entities, which are regulated by HIPAA (see below) or state laws.[97]

Health Insurance Portability and Accountability Act

The Health Insurance Portability and Accountability Act (HIPAA) of 1996 (correlative to certain provisions within the EU's GDPR) is a far-reaching federal law which protects patients' sensitive health data (any information related to health status, provision of healthcare, etc.) collected by a health care provider.[98] HIPAA regulates the sharing of protected health data and requires that health care providers ensure that data is protected against fraud and theft. In short, the goal of HIPAA is to limit the use of sensitive health data to those with a need to know. The Act covers all workers in covered entities, including those in healthcare facilities and private offices, students, non-patient care employees, health plans and insurance companies, billing companies, and electronic medical record companies, and applies to any healthcare information that might be linked to a specific patient. Entities must educate and train their workers in HIPAA compliance.

HIPAA's Privacy Rule provides individuals with certain rights, such as the right to access, the correction of inaccurate personal data, and the assurance that reasonable steps be taken to protect confidentiality. In addition, the Privacy Rule requires entities to keep individuals informed of protected data use and to document any disclosures. Protected data includes data of any size that is written, paper, spoken, and/or electronic, both within and outside a facility.

In order for protected health data to be used or disclosed, a patient must give signed consent. Legal exceptions (where professionals are not required to obtain consent before disclosure) include gunshot/stab wounds and other injuries sustained in crimes, child or elderly abuse, or infectious and reportable diseases, as well as in cases where legally required to disclose (such as with court orders). When information is disclosed, entities must make reasonable efforts to share only the minimum information necessary. HIPAA security officers must document and maintain security procedures, audit information systems, and implement security risk assessments. Violations of HIPAA may include stiff monetary penalties and potential imprisonment.

Transatlantic Data Privacy Framework

Post-*Schrems II*, the US and EU reached an agreement in March 2022 to replace the defunct Privacy Shield with a new Trans-Atlantic Data Privacy Framework.[99] Under the new framework, the US committed to strengthening its privacy safeguards over data transferred from the EU to the US, establishing a redress system, creating an Independent Data Protection Review court for EU citizens, and enhancing data protection oversight. It is hoped that this new framework will lead to more regulatory certainty and a freer flow of information between the two jurisdictions.

Proposed American Data and Privacy Protection Act

In June 2022, the US House of Representatives Committee on Energy and Commerce voted in favor of the bipartisan American Data and Privacy Protection Act (ADPPA).[100] This proposed legislation, if enacted, would harmonize US law with international privacy law frameworkers and serve as the US equivalent to the GDPR by creating a comprehensive framework for data privacy at the national level. The Act would federalize the protection of personal data, preempting state laws and creating a data privacy and security framework to protect the data of all Americans. In providing minimum privacy standards, states would be permitted to enact even stricter measures. With the goal of data minimization, ADPPA would allow any entities collecting, storing, and processing American consumer data to do so only if the purpose is specified in one of 17 permitted uses under the law. The Act provides for a private right of action for individuals, allowing them to pursue compensatory damages, injunctive relief, and attorney fees. The Federal Trade Commission (FTC) would be responsible for regulating and enforcing ADPPA.

GLOSSARY OF TERMS

AI: Artificial Intelligence

AIDS: Acquired Immunodeficiency Syndrome

ART: Antiretroviral Treatment

BD: Big Data

CDR: Call Detail Record

CIOMS: Council for International Organizations of Medical Sciences

CLM: Community-Led Monitoring

CT: Computerized Tomography

EMR: Electronic Medical Records

ERC: Ethics Review Committee

HIV: Human Immunodeficiency Virus

IRB: Institutional Review Boards

ML: Machine Learning

MNO: Mobile Network Operator

MRI: Magnetic Resonance Imaging

MSM: Men Who Have Sex with Men

NGO: Non-Governmental Organization

NLP: Natural Language Processing

PEPFAR: President's Emergency Plan for AIDS Relief

PH: Public Health (p. 25)

PHA: Public Health Authority

PLHIV: People Living with HIV

PrEP: Pre-exposure Prophylaxis

TasP: Treatment as Prevention

UN: United Nations

UNF's DIAL: United Nations Foundation's Digital Impact Alliance

VMMC: Voluntary Medical Male Circumcision

WHO: World Health Organization

ADPPA: American Data and Privacy Protection Act

AU: African Union

CoE: Council of Europe

DPA: Data Protection Act of 1998

DPIA: Data Protection Impact Assessments

DPO: Data Protection Officer

EUCJ: EU Court of Justice

FEC: Federal Executive Council

FIPP: Fair Information Practice Principles

FTC: Federal Trade Commission

FTCA: Federal Trade Commission Act

GDPR: General Data Protection Regulation

HIPPA: Health Insurance Portability and Accountability Act

ICCPR: International Covenant on Civil and Political Rights

IHR: International Health Regulations

NCSA: National Cyber Security Authority

NDPB: Nigerian Data Protection Bureau

NDPR: Nigeria's Data Protection Regulation

NITAU: National Information Technology Authority

NITDA: National Information Technology Development Agency

ODPC: Office of Privacy Commissioner

ENDNOTES

1 World Health Organization, "WHO Guidelines on Ethical Issues in Public Health Surveillance," Geneva: World Health Organization; 2017. License: CC BY-NC-SA 3.0 IGO.
2 World Health Organization, "Ethics and Governance of Artificial Intelligence for Health: WHO Guidance," Geneva: World Health Organization; 2021. License: CC BY-NC-SA 3.0 IGO
3 United Nations Development Group, "Data Privacy, Ethics and Protection: Guidance Note on Big Data for Achievement of the 2030 Agenda," resolution 45/95, November 2017, https://unsdg.un.org/resources/data-privacy-ethics-and-protection-guidance-note-big-data-achievement-2030-agenda.
4 See UNAIDS, "AIDSinfo," accessed June 1, 2023, https://aidsinfo.unaids.org/.
5 CDC Office of the Associate Director for Communication, "Eliminating HIV as a Global Public Health Threat," Centers for Disease Control and Prevention, last updated March 14, 2023.
6 Ann E. Kurth et al., "Combination HIV Prevention: Significance, Challenges, and Opportunities," Current HIV/AIDS Reports 8, no. 1 (March 2011): 62–72.
7 UNAIDS, "Fast-Track Targets: Understanding Fast-Track: Accelerating Action to End the AIDS Epidemic by 2030," June 2015.
8 Sherrie L. Kelly et al., "The Global Optima HIV Allocative Efficiency Model: Targeting Resources in Efforts to End AIDS," Lancet HIV 5, no. 4 (April 2018): 190–8.
9 Amitabh B. Suthar et al., "Towards Universal Voluntary HIV Testing and Counselling: A Systematic Review and Meta-Analysis of Community-Based Approaches," PLoS Medicine 10, no. 8 (August 2013): doi: 10.1371/journal.pmed.1001496.
10 Edinah Mudimu et al., "Individual and Community-Level Benefits of PrEP in Western Kenya and South Africa: Implications for Population Prioritization of PrEP Provision," PLOS One 15, no. 12 (2020) 15(12): https://doi.org/10.1371/journal.pone.0244761.
11 David J. Kedziora et al., "Optimal Allocation of HIV Resources among Geographical Regions, BMC Public Health 19 (2019): 1509.
12 K. Rivet Amico et al., "Adherence Support Approaches in Biomedical HIV Prevention Trials: Experiences, Insights and Future Directions from Four Multisite Prevention Trials," AIDS and Behavior 17, no. 6 (July 2013): 2143–55.
13 Amitabh B. Suthar et al., "Towards Universal Voluntary HIV Testing and Counselling: A Systematic Review and Meta-Analysis of Community-Based Approaches," PLoS Medicine 10, no. 8 (August 2013): doi: 10.1371/journal.pmed.1001496.
14 Edinah Mudimu et al., "Individual and Community-Level Benefits of PrEP in Western Kenya and South Africa: Implications for Population Prioritization of PrEP Provision," PLOS One 15, no. 12 (2020) 15(12): https://doi.org/10.1371/journal.pone.0244761.
15 Hana Kim et al., "Beyond HIV Prevalence: Identifying People Living with HIV within Underserved Areas in South Africa," BMJ Global Health 6, no. 4 (2021): doi 10.1136/bmjgh-2020-004089.
16 Mark Dworkin et al., "A Realistic Talking Human Embodied Agent Mobile Phone Intervention to Promote HIV Medication Adherence and Retention in Care in Young HIV-positive African American Men Who Have Sex with Men: Qualitative Study," JMIR Mhealth Uhealth 6, no. 7 (2018): doi: 10.2196/10211.
17 UNAIDS, "Key Populations," UNAIDS.org, 2023, https://www.unaids.org/en/topic/key-populations.
18 Carol S. Camlin and Edwin D. Charlebois, "Mobility and Its Effects on HIV Acquisition and Treatment Engagement: Recent Theoretical and Empirical Advances," Current HIV/AIDS Reports 16, no. 4 (August 2019): 314–23, https://doi.org/10.1007/s11904-019-00457-2.
19 Abby E. Rudolph et al., "Multiplex Relationships and HIV: Implications for Network-Based Interventions," AIDS and Behavior 21, no. 4 (April 2017): 1219–27, https://doi.org/10.1007/s10461-016-1454-2.
20 Susan J. Little et al., "Using HIV Networks to Inform Real Time Prevention Interventions," PLOS ONE 9, no. 6 (June 5, 2014): 1–8, https://doi.org/10.1371/journal.pone.0098443.
21 Mary Kate Grabowski et al., "Migration, Hotspots, and Dispersal of HIV Infection in Rakai, Uganda," Nature Communications 11, no. 1 (February 20, 2020): 1–12, https://doi.org/10.1038/s41467-020-14636-y.
22 Dylan Green et al., "Evidence of Sociodemographic Heterogeneity across the HIV Treatment Cascade and Progress towards 90-90-90 in Sub-Saharan Africa – a Systematic Review and Meta-analysis," Journal of the International AIDS Society 23, no. 3 (March 9, 2020): e25470, https://doi.org/10.1002/jia2.25470.
23 Ingrid T. Katz et al., "Impact of HIV-Related Stigma on Treatment Adherence: Systematic Review and Meta-Synthesis," Journal of the International AIDS Society 16, no. 3 Suppl 2 (November 13, 2013): 18640, https://doi.org/10.7448/IAS.16.3.18640.
24 Columbia University ICAP, "Population-based HIV Impact Assessment: Key Data to Guide the Global Response to the HIV Epidemic," Columbia University, accessed June 1, 2023: https://phia.icap.columbia.edu/.
25 Olga Tymejczyk et al., "HIV Treatment Eligibility Expansion and Timely Antiretroviral Treatment Initiation Following Enrollment in HIV Care," PLoS Medicine 15, no. 3 (March 23, 2018): doi: 10.1371/journal.pmed.1002534.
26 Oracle, "What Is Big Data?," June 1, 2023, https://www.oracle.com/big-data/what-is-big-data/.
27 Philippe Besse et al., "Can Everyday AI Be Ethical. Fairness of Machine Learning Algorithms" (arXiv, October 3, 2018), https://doi.org/10.48550/arXiv.1810.01729.
28 Isotta Landi et al., "Deep Representation Learning of Electronic Health Records to Unlock Patient Stratification at Scale," Npj Digital Medicine 3, no. 1 (July 17, 2020): 1–11, https://doi.org/10.1038/s41746-020-0301-z.
29 Laure Wynants et al., "Prediction Models for Diagnosis and Prognosis of COVID-19: Systematic Review and Critical Appraisal," The British Medical Journal 369, no. 1328 (March 31, 2020): 1–16.
30 Xianglong Xu et al., "Using Machine Learning Approaches to Predict Timely Clinic Attendance and the Uptake of HIV/STI Testing Post Clinic Reminder Messages," Scientific Reports 12, no. 1 (May 24, 2022): 8757, https://doi.org/10.1038/s41598-022-12033-7.
31 Anaelia Ovalle et al., "Leveraging Social Media Activity and Machine Learning for HIV and Substance Abuse Risk Assessment: Development and Validation Study," Journal of Medical Internet Research 23, no. 4 (April 26, 2021): e22042, https://doi.org/10.2196/22042.
32 Daniel J. Solove, Understanding Privacy, First Harvard University Press paperback edition (Cambridge, Massachusetts London, England: Harvard University Press, 2009). 161-165.
33 Daniel J. Solove, Understanding Privacy, First Harvard University Press paperback edition (Cambridge, Massachusetts London, England: Harvard University Press, 2009). 161-165.
34 Ari Ezra Waldman, "Social Theories of Privacy," in Privacy as Trust: Information Privacy for an Information Age, 1st ed. (Cambridge University Press, 2018), https://doi.org/10.1017/9781316888667.
35 Solon Barocas and Helen Nissenbaum, "Big Data's End Run around Anonymity and Consent," in Privacy, Big Data, and the Public Good, ed. Julia Lane et al., 1st ed. (Cambridge University Press, 2014),61, https://doi.org/10.1017/CBO9781107590205.004.
36 Effy Vayena and Lawrence Madoff, "Navigating the Ethics of Big Data in Public Health," in Oxford Handbook of Public Health Ethics, eds. Anna C. Mastroianni, Jeffrey P. Khan, Nancy E. Kass, Oxford University Press, 2019, 354-367.
37 Paul Ohm, "Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization," UCLA Law Review 57 (August 13, 2009): 1701–77.
38 Solon Barocas and Helen Nissenbaum, "Big Data's End Run around Anonymity and Consent," 52-56.
39 Luc Rocher, Julien M. Hendrickx, and Yves-Alexandre de Montjoye, "Estimating the Success of Re-Identifications in Incomplete Datasets Using Generative Models," Nature Communications 10, no. 1 (July 23, 2019): 3069, https://doi.org/10.1038/s41467-019-10933-3.
40 World Health Organization, "Ethics and Governance of Artificial Intelligence for Health: WHO Guidance," p. 41
41 https://arxiv.org/abs/cs/0610105
42 Lisa Singh et al., "Public Information Exposure Detection: Helping Users Understand Their Web Footprints," 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), Paris, France, 2015, pp. 153–61; doi: 10.1145/2808797.2809280.
43 Benjamin Y Anom, The ethical dilemma of mobile phone data monitoring during COVID-19: The case for South Korea and the United States, J Public Health Res 2022, Jul 5;11(3):22799036221102491. doi: 10.1177/22799036221102491.

44 Michelle Mello and C. Jason Wang, "Ethics and Governance for Digital Disease Surveillance," Science, May 29,200, vol. 368 issue 6494, 951-954.
45 Natasha Singer and Daisuke Wakabayashi, "Google to Store and Analyze Millions of Health Records," The New York Times, November 12, 2019. https://www.nytimes.com/2019/11/11/business/google-ascension-health-data.html.
46 Effy Vayena and Lawrence Madoff, "Navigating the Ethics of Big Data in Public Health," in Oxford Handbook of Public Health Ethics, eds. Anna C. Mastroianni, Jeffrey P. Khan, Nancy E. Kass, Oxford University Press, 2019, 354-367.
47 Linnet Taylor, "Safety in Numbers? Group Privacy and Big Data Analytics in the Developing World."
48 Daniel J. Solove, Understanding Privacy, 131.
49 Miranda Fricker, Epistemic Injustice: Power and the Ethics of Knowing (Oxford: Oxford University Press, 2007), and Kristie Dotson, "A Cautionary Tale: On Limiting Epistemic Oppression," Frontiers: A Journal of Women Studies 33, no. 1 (2012): 24–47.
50 United Nations Development Group, "Data Privacy, Ethics and Protection: Guidance Note on Big Data for Achievement of the 2030 Agenda."
51 Ariadne A. Nichol et al., "Diverse Experts' Perspectives on Ethical Issues of Using Machine Learning to Predict HIV/AIDS Risk in Sub-Saharan Africa: A Modified Delphi Study," BMJ Open 11, no. 7 (July 28, 2021): e052287, https://doi.org/10.1136/bmjopen-2021-052287.
52 World Health Organization, "Ethics and Governance of Artificial Intelligence for Health: WHO Guidance," June 28, 2021, 54.
53 Robert Sparrow and Joshua Hatherley, "The Promise and Perils of AI in Medicine," International Journal of Chinese & Comparative Philosophy of Medicine 17, no. 2 (January 1, 2019): 92, https://doi.org/10.24112/ijccpm.171678.
54 Ziad Obermeyer et al, "Dissecting racial bias in an algorithm to manage the health of populations," Science, 366, 447-453, 2019.
55 Milena Gianfrancesco et al, "Potential Biases in Machine Learning Algorithms Using Electronic Health Record Data," JAMA Internal Medicine, Nov. 2018, vol. 178, no. 11, 1544-1547
56 BMJ 2021; 372 doi: https://doi.org/10.1136/bmj.n304 (Published 16 March 2021)
57 David Leslie et al, "Does 'AI' Stand for Augmenting Inequality in the Era of Covid-19 Healthcare?" the BMJ 2021; 372:n304/ doi:10.1136/bmj.n304
58 Robert Sparrow and Joshua Hatherley, "The Promise and Perils of AI in Medicine," 94.
59 Brent Daniel Mittelstadt et al., "The Ethics of Algorithms: Mapping the Debate," Big Data & Society 3, no. 2 (December 1, 2016): 2053951716679679, https://doi.org/10.1177/2053951716679679. 6
60 World Health Organization, "WHO Guidelines on Ethical Issues in Public Health Surveillance;" United Nations Development Group, "Data Privacy, Ethics and Protection: Guidance Note on Big Data for Achievement of the 2030 Agenda."
61 World Health Organization, "WHO Guidelines on Ethical Issues in Public Health Surveillance," Geneva: World Health Organization; 2017. License: CC BY-NC-SA 3.0 IGO.
62 World Health Organization, "Ethics and Governance of Artificial Intelligence for Health: WHO Guidance,"
63 United Nations Development Group, "Data Privacy, Ethics and Protection: Guidance Note on Big Data for Achievement of the 2030 Agenda,"2017
64 Principles for Digital Development https://digitalprinciples.org
65 United Nations Conference on Trade and Development (UNCTAD, "Data Protection and Privacy Legislation Worldwide". https://unctad.org/page/data-protection-and-privacy-legislation-worldwide
66 Morrison Foerster, "Catch Up on Privacy Around the World on Data Privacy Day 2021!" https://www.mofo.com/resources/insights/210127-data-privacy-day
67 Coos, Andrada. "Data Protection Legislation Around the World in 2022". https://www.endpointprotector.com/blog/data-protection-legislation-around-the-world-in-2022/
68 United Nations. "Universal Declaration of Human Rights" (1948). https://www.un.org/en/about-us/universal-declaration-of-human-rights
69 United Nations. "International Covenant on Civil and Political Rights" (1966). https://www.ohchr.org/en/instruments-mechanisms/instruments/international-covenant-civil-and-political-rights
70 World Health Organization. "International Health Regulations" (2005). https://www.who.int/health-topics/international-health-regulations#tab=tab_1
71 World Health Organization. "International Health Regulations" (2008). https://www.who.int/publications/i/item/9789241580410
72 https://documents-dds-ny.un.org/doc/UNDOC/GEN/N14/707/03/PDF/N1470703.pdf?OpenElement; https://documents-dds-ny.un.org/doc/UNDOC/GEN/N13/449/47/PDF/N1344947.pdf?OpenElement
73 United Nations. "A/77/196: Principles underpinning privacy and the protection of personal data" (2022). https://www.ohchr.org/en/documents/thematic-reports/a77196-principles-underpinning-privacy-and-protection-personal-data
74 Council of Europe. Data Protection Website. https://www.coe.int/en/web/data-protection
75 Council of Europe. "Convention 108 - Convention for the protection of individuals with regard to the processing of personal data" (2018). https://rm.coe.int/convention-108-convention-for-the-protection-of-individuals-with-regar/16808b36f1
76 General Data Protection Regulation (GDPR). https://gdpr.eu/
77 General Data Protection Regulation (GDPR). "GDPR Article 5(1)." https://gdpr-info.eu/art-5-gdpr/.
78 Privacy Policies. "The 6 Privacy Principles of the GDPR" (2022). https://www.privacypolicies.com/blog/gdpr-privacy-principles/.
79 Privacy Policies. "GDPR: General Data Protection Regulation" (2022). https://www.privacypolicies.com/blog/gdpr/
80 General Data Protection Regulation (GDPR) "GDPR Article 5(2)". https://gdpr-info.eu/art-5-gdpr/.
81 Beverley Townsend (2022) The lawful sharing of health research data in South Africa and beyond, Information & Communications Technology Law, 31:1, 17-34, DOI: 10.1080/13600834.2021.1918905
82 European Commission. "European Data Governance Act" (2022). https://ec.europa.eu/digital-singlemarket/en/news/data-governance-act
83 Beverley Townsend (2022) The lawful sharing of health research data in South Africa and beyond, Information & Communications Technology Law, 31:1, 17-34, DOI: 10.1080/13600834.2021.1918905
84 "EU Court Expands Definition of Sensitive Data, Prompting LEgal Concerns for Companies". The Wall Street Journal. https://www.wsj.com/articles/eu-court-expands-definition-of-sensitive-data-prompting-legal-concerns-for-companies-11660123800
85 United Nations Conference on Trade and Development. "Data Protection and Privacy Legislation Worldwide". https://unctad.org/page/data-protection-and-privacy-legislation-worldwide
86 African Union. "African Union Convention on Cyber Security and Personal Data Protection" (2014). https://au.int/en/treaties/african-union-convention-cyber-security-and-personal-data-protection
87 Beverley Townsend (2022) The lawful sharing of health research data in South Africa and beyond, Information & Communications Technology Law, 31:1, 17-34, DOI: 10.1080/13600834.2021.1918905
88 Baker McKenzie. "Africa Data Security and Privacy Guide" (2021). https://www.bakermckenzie.com/-/media/files/insight/guides/2022/africa-data-privacy.pdf?sc_lang=en&hash=28FB5D9D701F5E4592625750A6F79A32
89 appKnox. "Recent Developments in Data Security Laws in Africa" (2021). https://www.appknox.com/blog/data-security-laws-in-africa
90 Baker McKenzie. "Data security and privacy laws develop across Africa" (2022). https://www.bakermckenzie.com/en/newsroom/2022/04/data-security-and-privacy-laws-across-africa
91 aker McKenzie. "How the European Union's General Data Protection Regulations influenced data privacy law in Africa" (2022). https://www.bakermckenzie.com/en/newsroom/2022/05/eu-general-data-protection-regulations
92 appKnox. "Recent Developments in Data Security Laws in Africa" (2021). https://www.appknox.com/blog/data-security-laws-in-africa
93 Staunton C, Tschigg K, Sherman G (2021) Data protection, data management, and data sharing: Stakeholder perspectives on

the protection of personal health information in South Africa. PLoS ONE 16(12): e0260341. https://doi.org/10.1371/journal.pone.0260341
94 Science. "A new law was supposed to protect South Africans' privacy. It may block important research instead" (2019). https://www.science.org/content/article/new-law-was-supposed-protect-south-africans-privacy-it-may-block-important-research
95 Federal Privacy Council. "Fair Information Practice Principles (FIPPs)." https://www.fpc.gov/resources/fipps/
96 Federal Trade Commission (FTC). "Federal Trade Commission Act Section 5: Unfair or Deceptive Acts or Practices". https://www.federalreserve.gov/boarddocs/supmanual/cch/200806/ftca.pdf
97 McGraw D, Mandl KD. Privacy protections to encourage use of health-relevant digital data in a learning health system. NPJ Digit Med. 2021 Jan 4;4(1):2. doi: 10.1038/s41746-020-00362-8. PMID: 33398052; PMCID: PMC7782585.
98 U.S. Department of Health and Human Services. "Health Information Privacy". https://www.hhs.gov/hipaa/index.html
99 The White House. "United States and European Commission Announce Trans-Atlantic Data Privacy Framework" (2022). https://www.whitehouse.gov/briefing-room/statements-releases/2022/03/25/fact-sheet-united-states-and-european-commission-announce-trans-atlantic-data-privacy-framework/
100 American Bar Association. "The American Data Privacy and Protection Act" (2022). https://www.americanbar.org/advocacy/governmental_legislative_work/publications/washington-letter/august-22-wl/data-privacy-0822wl/