

Intersectionality and LGBTI Data: Beyond Counting

How Data Agency Transforms LGBTI Data into Policy Impact

Edition October 2026

Acknowledgements

This report was written by Philip R. Crehan and Flavio Carrera Verdesoto, under the guidance of Dr. Elizabeth Lockwood, Francesca Perucci, and Raquel Santos Fontes. It is a collaboration between Open Data Watch and Koppa Lab. Many thanks go to the reviewers, Dr. Lee Badgett, Dr. Elizabeth Lockwood, Francesca Perucci, Raquel Santos Fontes, Caleb Orozco, and Eric Swanson. Special thanks are given to the people who shared their research and experiences with us:

Alex Garner, Managing Director, Mpact; Yuangi Mi, Data Analyst, Johns Hopkins University; Miltos Pavlou, Project Manager, Equality, Inclusion and Sustainability Unit, European Union Agency for Fundamental Rights (FRA); and three representatives from the National Institute of Statistics and Geography of Mexico (INEGI) – Octavio Heredia Hernández, Deputy Director General for Sociodemographic Surveys; Adriana Burgos Rojo, Director of Institutional Relations with the Private Sector and Non-Governmental Organizations; and Olivia Peralta, Deputy Director of State Engagement.

Terminology Note: LGBTI

We primarily use LGBTI terminology – referring to lesbian, gay, bisexual, transgender, and intersex people – as international organizations, civil society, and data stakeholders widely adopt it. At times, we use LGBTIQ to add “queer” to the acronym, often when citing source material that specifically included self-identifying queer people. We also use sexual orientation, gender identity and expression, and sex characteristics (SOGIESC) when referring to attributes rather than populations. Terminology in this field is context-specific and evolving, and alternative frameworks such as sexual and gender minorities or culturally grounded categories (for example, third gender) are used across different regions. Our choice of LGBTI and SOGIESC reflects their prevalence in global policy, statistical, and development discourse, as well as their usage among stakeholders represented in the case studies.

Executive Summary

This report explores how data agency enables LGBTI data to move through the Data Value Chain to influence public policy. The study is based on the Open Data Watch and Data2X's intersectionality in development data framework¹, which connects four dimensions: data production and use, data agency, linkages to policy analysis, and impact and monitoring. Applied across the Data Value Chain² (collection, publication, uptake, and impact), the framework treats better measurement not as an end in itself, but as part of a broader process through which communities can shape data, evidence can reach decision-makers, and public policy can produce more equitable outcomes.

Within this framework, the report focuses on data agency. Data agency is a core principle of the intersectional approach to data in the Open Data Watch and Data2X framework and refers to the ability and power of individuals and communities, especially those historically excluded or ignored by data systems, to shape how their data are collected, represented, governed, and used. Data agency is critical to translating data into action. While better measurement is important, data alone rarely change policy or improve outcomes. Evidence becomes influential when communities and institutions shape what is measured, how knowledge is interpreted, who has access to it, and how it connects to the institutions able to act on it.

The analysis is grounded in three comparative case studies representing complementary institutional approaches to LGBTI data: Mpact's community-led population size estimation of gay, bisexual, and other men who have sex with men; the European Union Agency for Fundamental Rights (FRA) recurring LGBTIQ Survey; and Mexico's National Institute of Statistics and Geography (INEGI) LGBTIQ+ survey. Together, these cases demonstrate how community organizations, regional institutions, and national statistical offices³ exercise different forms of data agency across the Data Value Chain.

Across the cases, the research finds that data agency plays a critical role in determining whether and how LGBTI data move through the Data Value Chain and generate public value. At the collection stage, agency shapes what becomes visible by influencing what is measured, whose experiences are represented, and what risks are acceptable in making populations visible. At the publication stage, participation, transparency, accessibility, and ethical safeguards help build trust and legitimacy while determining how data are shared and protected. At the uptake stage, agency helps connect evidence with communities, policymakers, statistical institutions, and other actors able to interpret and act on it. At the impact stage, sustained relationships and institutional engagement can translate evidence into changes in policies, programs, financing, statistical practices, and institutional capacity.

The cases also demonstrate that intersectionality is not achieved simply by collecting more variables or producing greater levels of disaggregation. It is

shaped by governance decisions about which identities and experiences can be measured credibly, meaningfully, ethically, and safely. INEGI incorporates multiple dimensions of identity while balancing community priorities against the practical constraints of a national survey. The FRA survey disaggregates sexual orientation, gender identity and expression, and sex characteristics while balancing intersectional analysis with comparability and sample-size limitations. Mpact uses digital data to estimate populations in settings where conventional approaches may be difficult or unsafe. However, its methodology limits which intersecting characteristics researchers can reliably observe. Intersectionality therefore emerges not as a single technical standard, but as a negotiated outcome shaped by communities, institutions, methodologies, resources, and political contexts.

Five cross-cutting findings also emerge. First, statistical visibility is created not by data alone, but by agency: communities, institutions, and partnerships help determine what becomes measurable and why. Second, sustained influence depends on recurring and institutionalized data production rather than one-off initiatives, allowing evidence, relationships, and institutional knowledge to build over time. Third, evidence gains authority and influence through trusted partnerships, particularly when community knowledge becomes connected with regional institutions, national statistical systems, and decision-makers. Fourth, meaningful innovation does not always depend on abundant resources: the case studies demonstrate how methodological choices, institutional mandates, community expertise, and partnerships can be combined to achieve influence even under resource constraints. And fifth, the impact of data can be twofold: evidence can contribute directly to changes in policies, programs, and resource allocation while also strengthening the relationships, technical capacities, and institutional practices needed to produce and use better LGBTI data over the long term.

The future of LGBTI data — and its potential to improve lives — cannot be assumed. In some settings, political and institutional developments are challenging the legitimacy of collecting, publishing, and using data on LGBTI people. Protecting the legal, institutional, and ethical conditions that allow communities to participate safely in data systems is therefore integral to strengthening intersectional data and ensuring that evidence can continue to inform public decisions.

The language of “counting” is an important starting point, but it can unintentionally portray LGBTI people as passive subjects of measurement. This report goes beyond counting by recognizing LGBTI communities as active agents in the Data Value Chain — not only people to be made visible in statistics, but participants in determining what is measured, how evidence is governed and interpreted, and how it is translated into policy. Ultimately, improving intersectional LGBTI data means producing better evidence as well as strengthening the agency, relationships, capacities, and policy linkages necessary for that evidence to contribute to more inclusive public policy and better outcomes for LGBTI people.

This report is intended for organizations working to strengthen inclusive data systems and evidence-based policymaking, including civil society organizations and LGBTI community groups, national statistical offices, researchers and academic institutions, the United Nations system, multilateral development banks, and donors.

Key Recommendations

- **Embed meaningful LGBTI participation and data agency throughout the Data Value Chain**, including clear roles for communities in shaping, governing, interpreting, and using data.
- **Design LGBTI data initiatives for uptake from the outset**, identifying intended users, policy pathways, decision-making processes, and opportunities to translate evidence into action.
- **Apply intersectional approaches throughout the Data Value Chain**, using participatory prioritization to balance community priorities with analytical rigor, respondent time and potential burden, sample size, comparability, and safety.
- **Strengthen LGBTI data ecosystems through sustained partnerships** among communities, civil society organizations, national statistical offices, researchers, regional and international organizations, and decision-makers, including institutions that can bridge community evidence with national systems.
- **Invest in statistical, analytical, and evidence-translation capacity**, particularly within LGBTI civil society and official statistical systems, while supporting methodological innovation for community members who remain difficult to reach.
- **Treat publication as an active stage of the Data Value Chain**, producing accessible and audience-specific knowledge products while enabling secure and appropriate access to underlying data.
- **Invest across the full Data Value Chain rather than in collection alone**, including multi-year measurement, publication, uptake, monitoring, evaluation, and evidence-to-policy translation.
- **Protect LGBTI data through ethical, privacy-preserving, and resilient governance**, including community ownership, secure access, data integrity, archiving, and contingency planning in politically sensitive contexts.
- **Promote learning, comparability, and accountability over time**, including cross-country knowledge exchange, appropriate harmonization, monitoring of data agency itself, and documentation of both successful and unsuccessful pathways from data to impact.

Part 1. Introduction

Intersectional data capture how a person's overlapping identities and characteristics interact to shape distinct and compounding experiences of advantage or disadvantage. Identities and characteristics can be wide-reaching, such as sex, sexual orientation, gender identity and expression, sex characteristics, age, race, ethnicity, disability,⁴ socioeconomic status, among others. Rather than treating each characteristic in isolation, intersectional data reveal patterns that single-attribute statistics obscure, while acknowledging the multiple dimensions of people's identities and experiences. This report explores the role of intersectional data along the Data Value Chain in reflecting the lived experiences of lesbian, gay, bisexual, transgender, and intersex (LGBTI) people and contributing to better policy outcomes.

Over the past two decades, investment in data on LGBTI people has expanded in many countries across civil society organizations, researchers, governments, national statistical offices, and international organizations. Important advances have expanded the evidence base through official statistics, community-led research, administrative systems, academic studies, and international development initiatives. Despite this progress, much of the resulting evidence remains fragmented across institutions and fails to become institutionalized knowledge capable of informing policy, financing, and institutional change. Valuable knowledge is frequently produced but not institutionalized, published but not widely used, or translated into advocacy without reaching the institutions capable of improving lives.

This disconnect reveals an important puzzle. The challenge is no longer that LGBTI people remain statistically invisible, although significant gaps in data and knowledge persist around the world.⁵ Rather, the challenge is understanding why some evidence moves beyond publication to policy while other evidence remains fragmented, overlooked, or unused. Producing more data is therefore only part of the solution. Equally important is understanding how existing evidence becomes connected, interpreted, and translated into meaningful public action.

Open Data Watch and Data2X's intersectionality and data framework provides a way to move intersectionality from an analytical concept into practical guidance for data systems. The concept note "Harnessing the Power of Intersectionality Data for Better Development Policies and Impact" organizes the work around four interconnected domains: the production of intersectional data, the agency of individuals and communities over how their data are governed and used, the link that translates data into policy, and the impact that policies informed by such data ultimately have on people's lives. Applied across the Data Value Chain, it treats better measurement not as an end but as one part of a larger process of shifting power in data governance toward more equitable outcomes.⁶

Data agency refers to the ability of individuals and communities, particularly those historically excluded or marginalized, to influence how data are designed, collected, interpreted, published, and ultimately used. Data agency is a key dimension of the intersectionality framework and is the mechanism through which translation of data into action can occur.⁷

While better measurement is essential, data alone rarely change policy or bring better outcomes. Evidence becomes influential when communities and institutions participate in shaping what is measured, how knowledge is interpreted, who has access to it, and how it informs public decisions. To examine this process, the report draws on Open Data Watch's Data Value Chain⁸ to examine how evidence moves from collection to publication, uptake, and impact.



Koppa Lab convenes its LGBTI Data Coalition in Mexico City, with non-governmental organizations, academics, UNFPA, and Mexico's National Statistical Office as part of its Data LIFT program (Data on LGBTI Inclusion for Transformation), which builds data agency among the coalition; June 2026

The analysis draws on three comparative case studies representing complementary governance models within LGBTI data ecosystems: Mpac's community-led population size estimation of gay, bisexual, and other men who have sex with men; the European Union Agency for Fundamental Rights recurring LGBTIQ Survey; and Mexico's National Institute of Statistics and Geography LGBTIQ+ surveys.

Together, these cases illustrate how community organizations, regional institutions, and national statistical offices exercise different forms of data agency across the four stages of the Data Value Chain. Rather than asking simply how LGBTI people can be counted, we ask how LGBTI data become influential, going beyond counting. By examining how data agency enables evidence to move through the Data Value Chain, we seek to contribute to stronger LGBTI data systems and broader efforts to build inclusive knowledge infrastructures in which those most affected by data have meaningful opportunities to shape the evidence. Ultimately, we ask a simple question: how does data agency enable LGBTI data to move through the Data Value Chain and influence public policy?

Part 2. A Brief Examination of LGBTI Evidence

The landscape of data on LGBTI people has expanded considerably over the past twenty years or so. Once limited to a few community surveys and advocacy reports, today's evidence base spans a diverse and growing ecosystem of data producers. These include legal and policy monitoring initiatives, public opinion surveys, academic research, official statistics, administrative records, civil registration systems, human rights observatories, private-sector and workplace data, market research, digital platform data, and community-generated studies designed to fill gaps left by official statistical systems. (See the taxonomy below.) Together, these efforts substantially increase what is known about the experiences of LGBTI people and create new opportunities for evidence-informed policymaking.

While the volume of evidence has grown considerably, the ecosystem itself remains fragmented, with limited coordination across sectors, inconsistent standards, and few mechanisms for connecting community-generated knowledge with official statistical systems or policy processes. As a result, valuable evidence often exists in parallel rather than in partnership, reducing its collective influence on public policy.

This fragmentation is evident across the ecosystem. Civil society organizations have pioneered participatory and citizen-generated data collection⁹, often reaching sub-groups and documenting experiences beyond the scope of official surveys. Academic institutions have contributed methodological innovation and rigorous analysis, while international organizations have supported regional and national surveys, composite indices, and comparative datasets. Administrative systems increasingly capture selected SOGIESC-related information in areas such as health, justice, and social protections, and a growing number of national statistical offices (NSOs) have incorporated SOGIESC variables into censuses, household surveys, labor force diagnostics, or dedicated LGBTI surveys.¹⁰ At the same time, newer forms of evidence such as digital trace data, commercial datasets, and online search engines offer additional ways to understand LGBTI people where conventional data remain unavailable.

LGBTI data availability remains constrained by several structural weaknesses. First, LGBTI data collection frequently develops through siloed approaches, with organizations producing valuable evidence that is rarely shared, harmonized, or connected to complementary initiatives. Second, limited statistical and analytical capacity prevents many organizations from fully analyzing and using the data they already possess. Third, many data producers lack proximity to power, the institutional relationships with governments, NSOs, development banks, and other decision-makers necessary to translate evidence into policy, financing, and systems change. Together, these factors create a persistent translation gap in which data are produced but not systematically activated for public decision-making, reinforcing a form of institutional invisibility.

LGBTI data sources encompass a growing range of data producers, methodologies, and institutional stakeholders. Collectively, these efforts provide an increasingly clear picture of the legal, social, economic, and lived experiences of LGBTI people. Yet they differ in their objectives, governance, and intended users. For analytical purposes, we group them into three broad domains.¹¹

Contextual data: what environments shape the lives of LGBTI people? These data describe the environments and contexts in which people live.

- This domain includes legal and policy data, for example whether governments criminalize LGBTI identities, or if there are anti-discrimination protections, family recognition, and broader legal frameworks. Public opinion and social attitudes data are also included, measuring social acceptance and norms. Additionally, data from human rights observatories are included that monitor violence, discrimination, and other rights violations. Finally, this domain includes composite indices and barometers that synthesize various indicators to benchmark and compare inclusion over countries and over time.

Individual data: what are the characteristics, experiences, and outcomes of LGBTI people? These data measure the characteristics and experiences of people themselves.

- This domain includes survey data, such as censuses, household surveys, labor force diagnostics, demographic and health surveys, and dedicated LGBTI surveys. Additionally, administrative data are generated through routine public services including health, education, justice, and social protections. Civil registration and vital statistics are also included here, where legal recognition of relationships or family status is documented by the state. Finally, this domain includes citizen data, often developed through participatory research by civil society organizations.

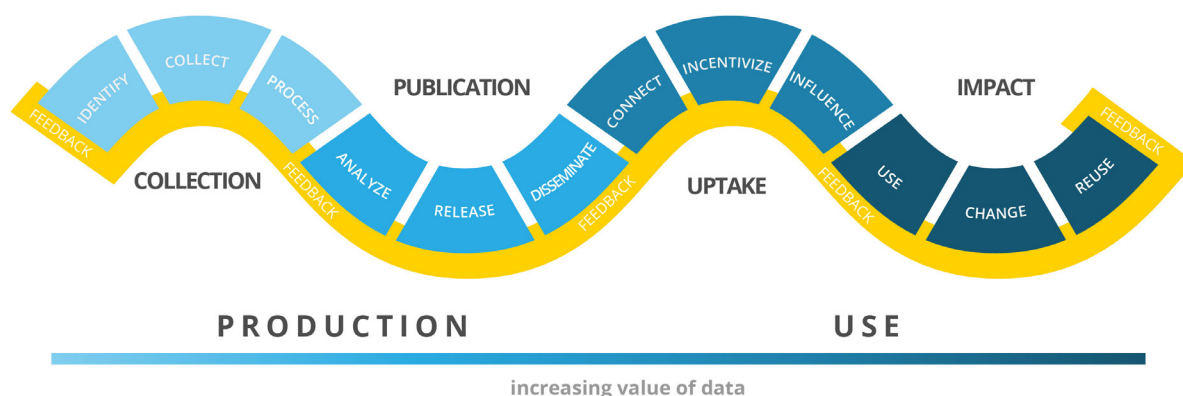
Knowledge and emerging data: What additional forms of evidence expand and deepen an understanding of LGBTI inclusion? These data expand the evidence base through research and nontraditional sources of information.

- This domain includes private-sector and company data, such as workplace inclusion surveys, employee resource group assessments, ESG reporting, or market research. Also here are financial transaction data, generated through banks or payment platforms that provide novel insights into the economic choices or financial inclusion of LGBTI people. Finally, this domain includes digital and big data drawn from online search behavior, social media, mobile applications, and other digital traces that generate insights.

Together, these domains demonstrate that the challenge facing LGBTI data is no longer simply one of scarcity. Severe data gaps persist across countries, sectors, and groups, including for lesbian and bisexual women, transgender people, and people born with intersex variations. At the same time, and in contexts with data, existing evidence is often fragmented, overlooked, or disconnected from decision-making. Individual initiatives have expanded across sectors and institutions, but often operate independently, using different methodologies, standards, and governance arrangements. The result is an expanding but only partially connected evidence base that has not yet developed into a coherent data system capable of informing policy and investments.

Part 3. Data Agency Across the Data Value Chain

Figure 1: The Data Value Chain



Source: Open Data Watch and Data2X, The Data Value Chain

The Data Value Chain

The Data Value Chain (DVC) provides a process-oriented perspective by tracing how data move from production to public value. Developed by Open Data Watch and Data2X, the DVC conceptualizes data not as an end in themselves but as part of a broader process through which evidence is transformed into public value. Rather than assuming that data automatically influence decisions once collected,

the framework recognizes that value is created through four successive stages, which provide the analytical framework for the remainder of our report:

- **Collection** examines how data are designed, governed, and generated, including the role of community participation in determining what is measured and how.
- **Publication** considers how data are analyzed, disseminated, and made accessible, as well as the institutional and political factors that influence visibility, transparency, and trust.
- **Uptake** focuses on how evidence is interpreted, mobilized, and incorporated into advocacy, government decision-making, and institutional processes.
- **Impact** examines the extent to which data contribute to sustained changes in policy, programs, financing, and broader institutional systems.

Intersectionality is treated as a cross-cutting principle within the Data Value Chain. Intersectionality recognizes that multiple, overlapping identities within systems of power shape experiences of exclusion rather than any single characteristic alone. For LGBTI people, sexual orientation, gender identity and expression, and sex characteristics (SOGIESC) intersect with race, ethnicity, disability, age, migration status, geography, socioeconomic status, and other dimensions to produce diverse experiences and outcomes. Throughout the DVC, intersectionality helps shape which data are collected, who is represented, how evidence is communicated and used, and how it can influence public policy. Throughout the report, we apply an intersectional lens, while recognizing that the degree of intersectional coverage varies across case studies.

“Agency is the ability to act and affect changes in spheres that are important to the individual or the collective [...] Data agency in intersectionality includes the basic tenet that counting the excluded and making them visible in statistics acknowledges their status as citizens with rights and responsibilities. But it goes beyond this basic principle and calls for excluded populations to have a say in, control or ownership over the production and use of data”

– Open Data Watch & Data2X, Harnessing the Power of Intersectionality
Data for Better Development Policies and Impact: A Concept Note

LGBTI Data Agency

It is within emerging LGBTI data ecosystems that data agency across the Data Value Chain becomes consequential. As these ecosystems expand, the central question is no longer only who collects data, but shifts to who shapes it, governs it, interprets it, and ultimately determines how data are used.

In the intersectionality framework, data agency is the mechanism through which that movement occurs. Rather than viewing the DVC as a purely technical sequence, we examine it as a governance process in which decisions are continuously made about what data are collected, whose knowledge is represented, how evidence is communicated, who can use it, and whether it ultimately informs policy and institutional change. At every stage, opportunities exist for LGBTI stakeholders to shape (or be excluded from) these decisions. In this sense, data agency provides the connective mechanism through which fragmented sources of evidence can move across the DVC—from collection to publication, uptake, and impact. We examine how community participation, institutional relationships, and governance influence the translation of evidence into policy and financing.

Our analysis finds that data agency is distributed across multiple stakeholder groups. While the case studies primarily examine agency exercised by LGBTI communities and organizations themselves, they also demonstrate the important roles of regional organizations, national statistical offices, and researchers. Agency is neither static nor concentrated within a single entity or group, rather, it shifts across the stages of the DVC.

Data agency does not imply that more data should always be collected or used throughout the DVC. For LGBTI people living under criminalization, authoritarian rule, or who face political targeting, data collection and usage may expose them to danger or other challenges. In those contexts, an expression of agency may be the deciding not to collect certain data, or to limit its publication, or to retain community control over their uptake and use. Safety, ethics, confidentiality, informed participation, and community-led standards are of the utmost importance.¹²

Methodology

We conducted a targeted desk review of academic literature, technical reports, and grey literature on LGBTI data and data agency. Our review included publications from civil society organizations, international organizations, national statistical offices, and academic sources, with particular attention to examples in which LGBTI data informed policy, institutional practice, or other tangible outcomes. This review also informed the selection of case studies.

Three case studies were selected to provide geographic and institutional diversity. Collectively, the cases span five regions and dozens of countries through the

initiatives they lead, and represent a community-based organization, a regional organization, and a national statistical office. Individually, each study was selected based on its demonstrated engagement with data, participation of LGBTI communities in data processes, and evidence of policy, programmatic, or institutional impact. In addition to interviews, we reviewed relevant reports, survey instruments, technical documentation, methodological papers, websites, and other public outputs associated with each study.

Semi-structured interviews were conducted using virtual meeting platforms with six key informants representing the three case studies. Interviews were recorded with participants' consent and transcribed using artificial intelligence. The interview guide was adapted from the "Intersectional Data and Persons with Disabilities"¹³ report and revised to examine how data agency can operate within LGBTI data systems. Interview transcripts were analyzed using a deductive thematic coding approach based on the interview guide and the stages of the DVC, while allowing additional themes to emerge inductively from responses. Across all three cases, a cross-cutting analytical lens examined where and how LGBTI community agency was strongest and weakest throughout the DVC. Interview findings were triangulated with documentary evidence to verify information, strengthen the analysis, and provide additional detail.

Rather than seeking statistical representativeness, we adopt a comparative case study design to examine how data agency operates across institutional settings. These case studies are illustrative and are intended to demonstrate different pathways through which data agency operates across the Data Value Chain and, therefore, cannot represent the full diversity of all LGBTI data initiatives. Because the cases were selected partly for demonstrated impact, the analysis identifies pathways and enabling mechanisms rather than establishing the causal effect of agency.

Case Studies

The three case studies anchor our comparative analysis. Representing a community-based organization, a regional organization, and a national statistical office, they examine how different governance models influence the collection, publication, uptake, and impact of LGBTI data. Comparing these models reveals common barriers and enabling conditions that shape whether evidence ultimately informs policy and institutional change. Across the three cases, the central analytical question is: how does data agency enable LGBTI data to move through the Data Value Chain and ultimately influence policy?

- 1. Mpact's analysis in Africa, Europe, Asia, Latin America, and North America.** Mpact is a community-based organization for gay men's health and rights that works with key populations (gay and bisexual men, and other men who have sex with men, or "GBMSM"¹⁴) in the context of the HIV epidemic.¹⁵ In 2018 and 2026, Mpact supported and led new population

size estimates (PSE) of GBMSM using digital data (social media, dating apps, and search engines). Within HIV programming, official estimates generated by the government and used by the Joint United Nations Programme on HIV/AIDS (UNAIDS) are linked to funding and the provision of services yet are widely seen as undercounted. The 2026 PSE was implemented through Mpact's program "Queer Data" in partnership with Johns Hopkins University and LGBTI community-based organizations in Senegal, Ghana, Mozambique, and Indonesia. The program seeks to provide more realistic estimates of GBMSM for local partners to use in their advocacy and to receive more funding and better services from their governments.

- 2. The European Union's Fundamental Rights Agency survey in 30 countries.** The FRA is an independent agency within the European Union (EU) that produces evidence-based guidance on fundamental rights across the EU member states and candidate countries. Its LGBTIQ¹⁶ survey began with a request from the European Parliament between 2010–2011, and had three iterations in 2012, 2019, and 2023. The latest iteration of the survey represents one of the largest datasets of LGBTIQ people in the world (with over 100,500 respondents), surpassed only by the FRA's previous survey in 2019. This survey and its findings provide an evidence base for the EU's LGBTIQ Equality Strategies, among other policies and standards.¹⁷
- 3. Mexico's NSO, INEGI, and its national LGBTI data collection efforts.** The Encuesta Nacional sobre Diversidad Sexual y de Género (ENDISEG) is the first national survey focused on LGBTIQ+ people in Mexico.¹⁸ The ENDISEG is comprised of two research efforts: an in-person nationally representative survey (around 44,200 respondents) and an online, non-probabilistic survey (called "ENDISEG Web") with 14,364 respondents, of whom 7,060 identified as LGBTIQ+.¹⁹

Collection: How Agency Shapes What Is Measured

The case studies show that data agency begins before data are collected. Agency starts with the power to define the problem, shape the research question, choose the instrument, and decide which identities and experiences are made statistically visible. Across the three case studies, collection is not just a technical stage, but the first governance decision in the Data Value Chain: who gets to decide what counts, how people are counted, and what risks are acceptable in making a population visible.

The studies show three different models of agency. In Mexico, agency emerged through pressure from civil society and was then institutionalized through INEGI's official statistical mandate. Those demands moved through government institutions and became a dedicated national survey rather than a few questions added to the census, culminating in the ENDISEG projects.

“It wasn’t that the Institute recognized that lack of visibility on its own. There were many social movements in which collectives and people from this sector were calling out that absence of information.”

— Interview with Octavio Heredia Hernández, INEGI

The FRA case shows a different model: a regional institution using its legal and policy mandate to create a recurring, comparable survey, while drawing on civil society participation to shape the instrument. In the interview with the FRA, the interviewee said rights holders were involved at every stage including consultation, translation checks, questionnaire testing, and dissemination through a network of around 800 organizations across 30 countries. In the end, the FRA retained final authority over methodology and questions asked, but did so by balancing community input, EU policy needs, and analytical rigor. This illustrates a central tension in data agency: participation does not always mean full control, but it can still meaningfully shape what is measured.

Mpact represents the most community-led model, with a non-governmental organization leading the analysis and community-based organizations using the evidence to advance their own health needs. The PSE responds to settings where official estimates of GBMSM are widely recognized as incomplete, yet conventional survey methods may be infeasible because of stigma, criminalization, or weak political commitment. Rather than relying solely on traditional epidemiological approaches, Mpact and its research partners developed an innovative methodology combining digital trace data from online platforms with community expertise to estimate populations that official systems often fail to measure. As a 2018 publication stated, the PSE is "a method leveraging different categories of same-sex interests on Facebook, combined with a specific gay-oriented app (Hornet), demonstrating significantly higher estimates than those officially reported."²⁰

But this form of collection also has limits: the PSE can capture geography and behavior more readily than identity, meaning intersectional analysis is constrained by the method itself. For Mpact, this method allowed them to capture digital proxies of same-sex interest, behavior, and platform use, but not gender identity and expression, meaning the PSE can only estimate same-sex attracted men but not transgender women.

“Big data tells you where you are and what you’ve done [...] geography and behavior. Everything else becomes an assumption, and you have to approach it differently.”

— Interview with Alex Garner, Mpact

Across the three case studies, agency shapes data collection in three main ways. First, agency determines whether invisibility becomes a statistical problem: civil society pressure in Mexico, EU policy demand in Europe, and community advocacy in HIV programming all transformed absence into a mandate for measurement. Second, agency shapes the instrument itself: ENDISEG prioritized a dedicated, privacy-protecting survey; the FRA expanded SOGIESC measurement across three dimensions (away from a one-dimensional box); and Mpact adapted digital methods to oppressive and undercounted contexts. Third, agency defines the ethical boundaries of visibility. All three cases show that counting LGBTI people can create policy value, but also requires safeguards around privacy, confidentiality, stigma, and political risk.

The lesson for collection becomes apparent: data agency is not an optional participatory add-on. Agency is what makes collection relevant, legitimate, and safe. Without agency, data systems may count the wrong things, ask questions in ways that communities mistrust, or produce evidence that cannot be used. With agency, collection becomes the first step in moving LGBTI data through the DVC toward public policy.

Publication: How Agency Shapes Visibility, Trust, and Access

The publication stage is more than releasing findings or datasets. It is the process by which evidence becomes visible, builds legitimacy, and becomes available for use. Data agency extends beyond collection to encompass decisions about how results are analyzed, communicated, protected, and shared. Across the three case studies, publication reflects different institutional norms, yet all grapple with the same challenge: maximizing access and policy relevance while safeguarding the people represented in the data.

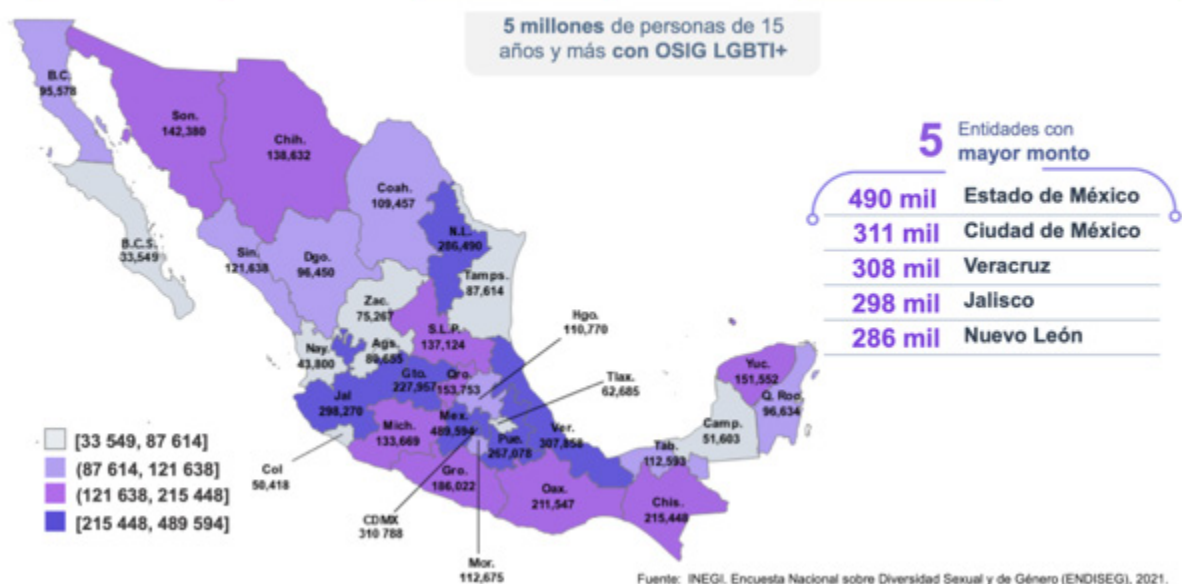
Mexico's ENDISEG illustrates publication as a public good anchored in official statistics. Consistent with INEGI's principles of transparency, the survey was accompanied by a comprehensive set of outputs, including methodological notes, the questionnaires, statistical tables, anonymized microdata, and public presentations. With this publication stage, INEGI sought to maintain strict confidentiality protections for respondents. Access to sensitive information is governed through secure microdata laboratories, balancing openness with legal obligations under Mexico's statistical system.

“INEGI publishes practically everything, except information that could violate the principle of confidentiality. That is a milestone for the Institute in terms of transparency of information.”

— Interview with Octavio Heredia Hernández, INEGI

Figure 2: Estimated population with diverse sexual orientations and gender identities in Mexico

TOTAL OSIG LGBTI+ POR ENTIDAD



Maps the estimated 5 million people aged 15 and older who identify as LGBTI+ or have diverse sexual orientations and gender identities (SOGI; in Spanish, OSIG) across Mexico's 32 states, ranked by the five states with the highest totals.

Source: INEGI, ENDISEG (2021)

The FRA adopts a similarly open approach but frames publication around accessibility and policy translation. Rather than producing a single technical report, it releases multiple products tailored to different audiences, including a flagship report, interactive Data Explorer, country fact sheets, and anonymized open microdata.²¹ At the same time, publication is limited to protect respondents, with potentially identifying information removed and small population groups protected from re-identification. The FRA's experience demonstrates that openness and confidentiality are complementary rather than competing objectives.

By contrast, Mpact conceives publication as an extension of community ownership. The organization follows a dual dissemination strategy: (1) academic publication to establish methodological credibility and (2) advocacy-oriented products designed for local organizations that want to influence resource allocation. Publication therefore becomes an act of returning knowledge to the communities of focus, rather than simply communicating research findings.

“All of this research, all of this work belongs to the community—it doesn’t belong to an institution... I’m very much about making it as accessible to anyone who wants it.”

— Interview with Alex Garner, Mpact

Across the three case studies, publication emerges as a key site of data agency, serving three interconnected functions. First, agency transforms data into public visibility, determining which experiences become part of institutional knowledge and public debate. Second, agency establishes credibility and trust through methodological transparency, appropriate safeguards, and products that different audiences can understand and use. Finally, agency governs access, balancing the benefits of open data with the ethical obligation to protect respondents from harm. While the mechanisms differ (for example, official statistical release protocols versus regional open-data platforms versus community-centered dissemination) the underlying principle is consistent: publication is an exercise of data agency that determines whether evidence remains confined or becomes a resource capable of influencing public policy.

Intersectionality also shapes what publication can reveal. The FRA’s Data Explorer lets users see 76 SOGIESC combinations, but its public outputs deliberately keep the information on race and ethnicity minimal to protect small subgroups from re-identification, according to the interview. Similarly, the interviews revealed that INEGI’s published tables extend disaggregation to disability, Indigenous and Afro-descendant identity, and skin tone alongside SOGIESC, where sample sizes allow. Mpact’s estimates cannot be disaggregated by gender identity at all, because the method only measures geography and proxies for same-sex attracted behavior. In each case, publication choices determine which intersecting subgroups become visible and which remain aggregated.

The cases also reveal that publication is not the endpoint of the Data Value Chain. Rather, publication is the bridge between evidence and action. Data that are technical and rigorous but remain inaccessible or incomprehensible to communities are unlikely to influence decision-making. Effective publication

therefore requires more than transparency; it requires deliberate choices about audience, communication, legitimacy, and stewardship that enable evidence to move into the next stage of the Data Value Chain.

Uptake: How Agency Translates Evidence into Decision-Making

Uptake is the stage at which evidence moves beyond publication and begins informing decisions, advocacy, programming, and institutional action. While collection determines what is measured and publication determines what becomes visible, uptake determines who is able to act on the evidence and for what purpose. Across the three case studies, data agency is expressed not simply through access to data, but through the capacity and legitimacy to translate evidence into decisions.

The three governance models demonstrate complementary pathways of uptake. INEGI's ENDISEG entered government institutions through the legitimacy of official statistics, providing ministries, public agencies, researchers, and civil society with a common evidence base for understanding the challenges affecting LGBTI people. Because the survey was produced within the national statistical system, its findings carried institutional authority that facilitated use across government while remaining publicly available.

The FRA illustrates uptake at a regional scale. Rather than expecting a single report to drive change, the agency deliberately created an ecosystem of users. EU institutions, member states, academics, journalists, and civil society organizations all engage with the LGBTIQ findings through different products, while organizations such as ILGA-Europe and academics extend the analysis using the openly available microdata. Country fact sheets become tools for constructive conversations with governments, translating evidence into dialogue rather than confrontation. As the FRA observed, there has been tremendous interest in the survey findings.

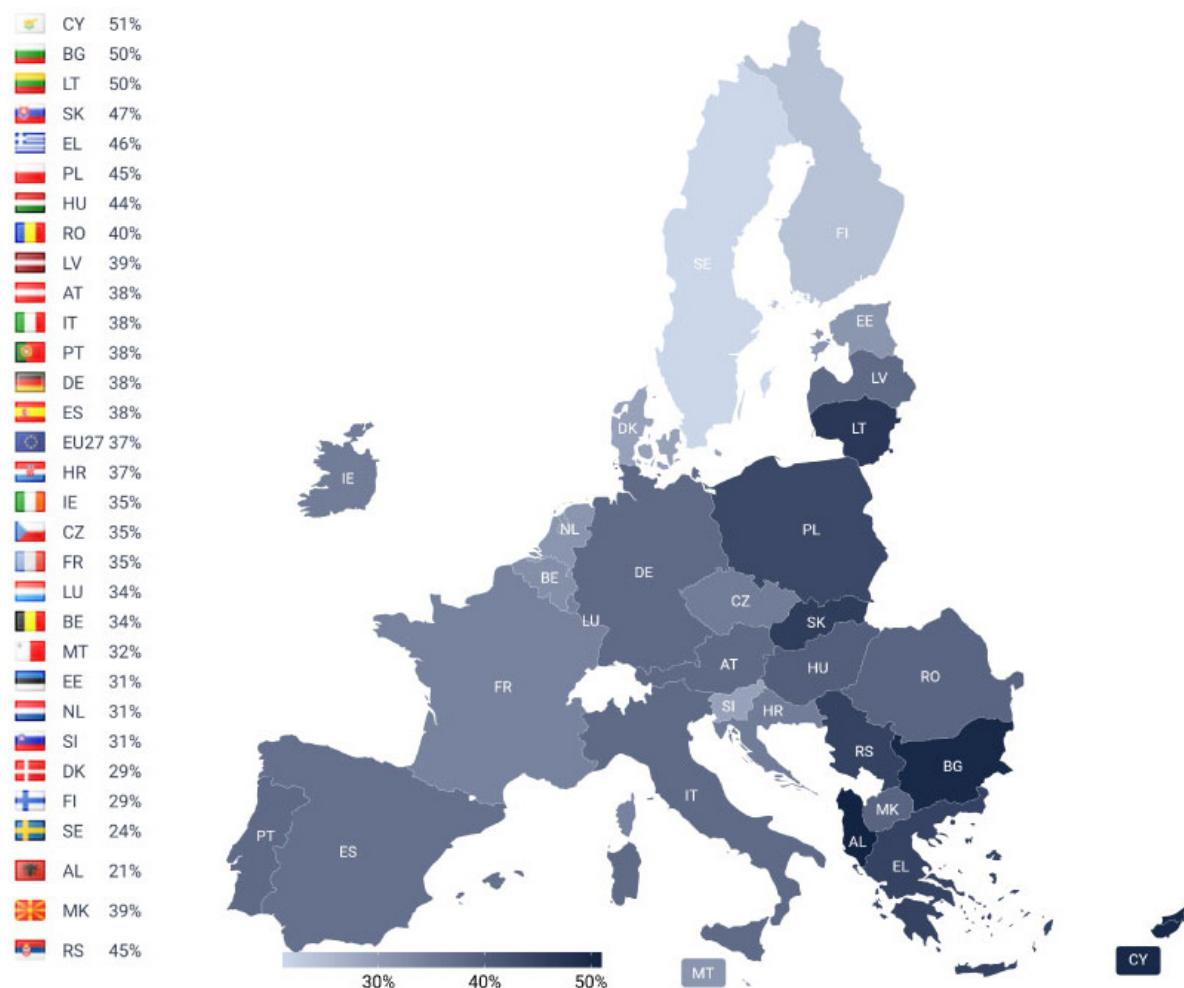
“Among all the data explorers across FRA surveys, when we looked internally to the FRA survey data explorers traffic, we found that the LGBTIQ one had gathered more than 70% of the traffic last year.”

— Interview with Miltos Pavlou, FRA

Figure 3: Perceived discrimination against LGBTI people across Europe

In the past 12 months have you personally felt discriminated against due to being LGBTI in any of 8 areas of life?

That answered Yes



© FRA - All rights reserved - EU LGBTIQ Survey III, 2023 | Discrimination in areas of life > Felt discriminated in the 12 months before the survey in any of 8 areas of life - D1_1 | Answer: Yes

FRA
EUROPEAN UNION AGENCY
FOR FUNDAMENTAL RIGHTS

Visualization from the question, "In the past 12 months, have you personally felt discriminated against due to being LGBTI in any of eight areas of life?"

Source: FRA, LGBTI Data Explorer (2023).

Further, the FRA LGBTIQ survey's influence also depended on demonstrating methodological robustness. Although the survey was a large-scale non-probability online survey rather than population-based, stratified recruitment targets and post-survey weighting were used to improve sample representativeness, while extensive dissemination broadened participation and geographic reach. As the technical report notes, "the large and balanced sample ensured that data processing and analysis could make the results as representative as possible. [...] Remaining imbalances were addressed by applying correction techniques, such as weighting."²²

Mpact represents a different pathway. For the organization, the primary users are not governments but on-the-ground community advocates. Population estimates are designed to strengthen negotiations with donors, ministries of health, UNAIDS, and financing mechanisms by providing communities with credible evidence to support claims for greater resources. Technical expertise remains largely with Mpact's research partners, especially universities, while local organizations are equipped to use the findings strategically.

"It needs to be a tool that people can use to improve their advocacy or make it more effective—and that advocacy is around resources, policy change, representation, visibility."

— Interview with Alex Garner, Mpact

Across the three cases, uptake depends on more than making data available. Uptake requires institutional credibility, analytical capacity, and proximity²³ to decision-makers. The community-led model builds negotiating power from the ground-up. The EU regional model mobilizes evidence through networks and transnational institutions. And the national statistical model embeds evidence directly within government systems. Together, they demonstrate that data agency extends beyond data ownership to encompass the ability to interpret, communicate, and mobilize evidence within the institutions where decisions are made.

The cases also reveal that the principal barriers to uptake are not necessarily technical. Rather, they stem from limited analytical capacity, weak institutional relationships, and insufficient pathways connecting evidence producers with policymakers. These findings reinforce the argument introduced earlier: fragmentation exists not only across LGBTI data initiatives, but also in the systems that translate data into action. Data agency helps bridge that gap by connecting evidence with the stakeholders capable of transforming it into public decisions.

Uptake is also shaped by how much intersectional detail is available upon which to act. The FRA's open microdata has allowed ILGA-Europe (a large regional non-governmental organization, or NGO) to lead its "Intersections 2.0" analyses²⁴ of health outcomes among sub-groups, and academic researchers have used the data to do their own analyses beyond the FRA's cross-tabulations. Similarly, INEGI's disability and Indigenous-identity variables allow other parts of the government and researchers to target policy questions narrower than SOGIESC alone.

Impact: How Agency Produces Institutional Change

Impact represents the final stage of the Data Value Chain, where evidence contributes to lasting changes in policies, institutions, financing, and community outcomes. Causal links are rarely straightforward between research and outcomes because evidence typically must be paired with political leadership, advocacy, legal developments, and broader social acceptance. Nevertheless, each case study still demonstrates that data agency can increase the likelihood that evidence will influence change by ensuring that data remain relevant, credible, and actively mobilized throughout the policy process.

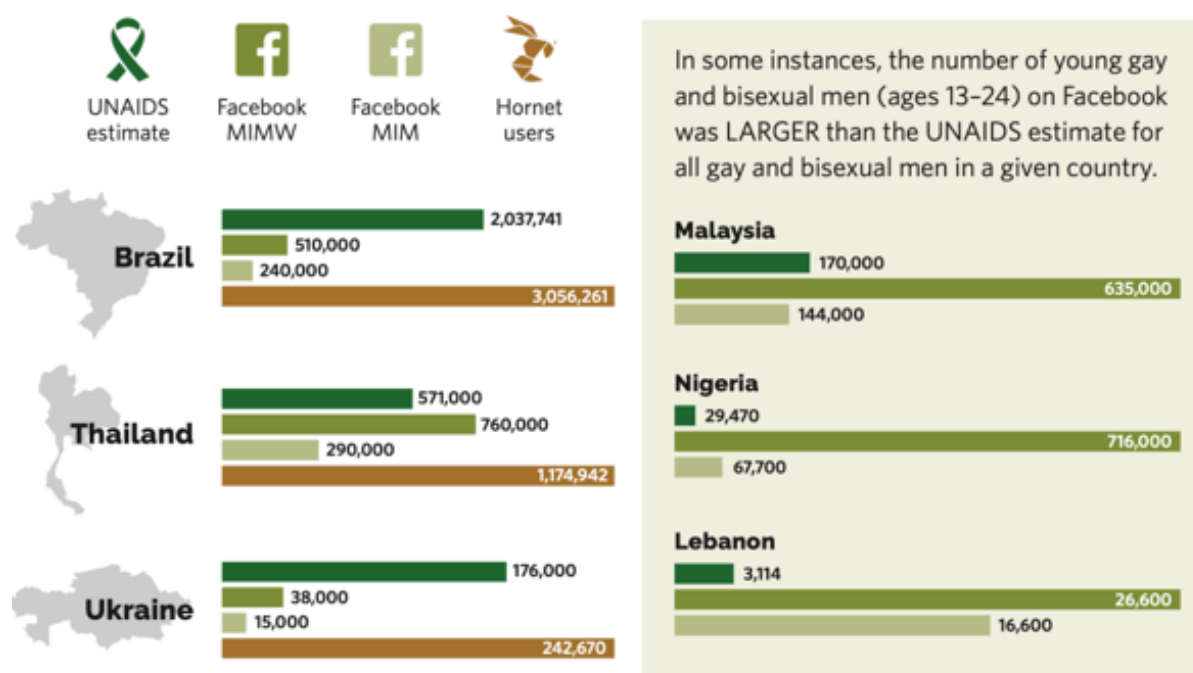
The three case studies reveal three separate pathways to impact. For Mpact, the impact is measured by whether community-generated evidence changes, or is adopted alongside, official population estimates, and how those new estimates can be used to influence HIV services. Estimates from the first PSE were cited in the UNAIDS Key Populations Atlas for more than a dozen countries, according to the interview with Mpact. The American Foundation for AIDS Research (amfAR) also cited those estimates in planning documents for PEPFAR, the U.S. HIV funding program.²⁵ Partly because of the use by UNAIDS and amfAR, this allowed select non-governmental organizations on-the-ground to advocate for more resources from their governments. Similarly, Mpact defines impact more broadly as a process of strengthening community capacity to advocate for itself and ensuring that methods remain in community hands even as global HIV institutions are shut down and as digital data sources shrink.

Additionally, the Mpact study highlights the broader policy consequences of inadequate population size estimates. HIV financing is linked to official UNAIDS population estimates, meaning that inaccurate or implausibly low estimates can have significant downstream effects on resource allocation and service provision. In many countries, the absence of reliable data reflects the same structural barriers that increase risk to acquire HIV, including criminalization, stigma, and limited political will. As the 2018 report observed, “we know the least about the numbers and social experiences of the most vulnerable communities of GBMSM in settings deemed the most hostile.”²⁶ That report further suggests that financing systems may create unintended incentives to maintain low estimates: “[...] high HIV service coverage goals may disincentivize accurate size estimates for countries with implausibly low or absent population estimates.”²⁷ In this way, methodological weaknesses can become embedded within financing systems, reinforcing statistical invisibility and perpetuating the under-allocation of resources to the populations who need it the most.

Figure 4: Mpart's estimates of gay and bisexual men in countries

What We Found:

UNAIDS estimates for the total number of gay and bisexual men in each country were **almost always smaller** than gay and bisexual Hornet or Facebook users.



Estimates of the number of gay and bisexual men in each country were almost always lower than estimates based on gay and bisexual Hornet or Facebook users.

Source: Mpart, HIV Indicators for Gay, Bisexual and Other Men Who Have Sex with Men: A Technical Brief (2022), based on data from Baral et al. (2018).

The FRA demonstrates institutional impact at a regional level. According to the interview conducted for the case study, the survey provides the evidence base for successive EU LGBTIQ Equality Strategies, informs work by the European Parliament and European Commission, contributes to Council of Europe recommendations, and increasingly influences national statistical systems by including SOGIESC measures. Rather than positioning itself as an advocacy organization, the FRA emphasizes constructive engagement with governments through evidence.

“We’re not an academic center of research; we provide data about what the reality is... We are there as allies, trying to show them how they can improve—not to finger-point, to shame and blame, because the result might already do that.”

— Interview with Miltos Pavlou, FRA

INEGI illustrates institutionalization within an official statistical system. One of the strongest impacts is the establishment of data on LGBTI people as a legitimate component of Mexico's statistical architecture. Through sustained collaboration with civil society and methodological innovation, ENDISEG transformed what had previously been absent from official statistics into an evidence base that future governments, researchers, and civil society organizations can revisit, expand, and use.

Taken together, the three cases suggest that data agency produces impact through different institutional pathways but similar underlying mechanisms. For Mpact, community leadership generates legitimacy and advocacy power. For the FRA, regional institutions translate evidence into coordinated policy frameworks. And for INEGI, the NSO institutionalizes measurement within official systems. In each case, impact depends not only on data quality but on sustained relationships among communities, researchers, statistical institutions, and policymakers.

The impact of applying intersectionality to data

The case studies highlight that intersectionality shapes whose outcomes change. For example, the FRA's data formed part of the evidence base for the Council of Europe recommendation specifically on the rights of intersex people²⁸, illustrating how measuring sex characteristics separately can make intersex-specific inequalities visible to policymakers. INEGI's disability and Indigenous-identity variables similarly open the door to policy attention on LGBTIQ+ people who have a disability or are Indigenous, though INEGI treats this as an early frontier.

Across the four stages of the Data Value Chain, the case studies show that intersectionality is not just measurement but a question of governance. In the case studies, intersectionality did not emerge simply by adding variables but rather depended on governance decisions about which identity aspects could be measured credibly and safely. Intersectionality was therefore unevenly applied across the case studies. INEGI incorporated multiple dimensions of identity through its national surveys, balancing community priorities against questionnaire constraints. The FRA survey can disaggregate SOGIESC into

separate dimensions (SO, GIE, SC), enabling many identity combinations while preserving comparability across countries. But the FRA also acknowledged limitations in capturing smaller population groups and additional identity characteristics. Mpact faced an opposite challenge: although digital population estimation effectively measures geography and (proxies of same-sex attracted) behavior, it cannot reliably capture other intersecting identities without introducing assumptions. Altogether, these cases suggest that intersectionality is not a single standard, but a negotiated governance outcome shaped by many stakeholders and conditions throughout the Data Value Chain.

Part 4. Cross-Cutting Key Findings and Recommendations

The preceding analysis of collection, publication, uptake, and impact across Mpact, the FRA, and INEGI demonstrates that data agency is not a peripheral consideration in LGBTI data systems. Rather, agency can determine whether data translate into public value. Five cross-cutting findings also emerge from the case study comparative analysis, particularly pertinent to the Data Value Chain and LGBTI data.

Cross-Cutting Key Findings

Visibility. Across all three case studies, statistical visibility emerged as the first and most fundamental contribution of LGBTI data. ENDISEG produced Mexico's first nationally representative official evidence on LGBTI people; the FRA created the largest comparative dataset on LGBTIQ experiences across Europe; and Mpact generated credible estimates of GBMSM populations in countries where official statistics frequently undercount them. Yet the cases also reveal that statistical visibility rarely emerges through institutional initiatives alone. Rather, invisibility often became a recognized statistical problem when communities, advocates, or policymakers possessed sufficient agency to demand measurement. Civil society organizing in Mexico, policy demand from European institutions, and community advocacy within HIV programming each transformed absence into a mandate for data collection. Statistical visibility can also reinforce broader social visibility, helping communities make their experiences, inequalities, and policy needs more tangible.

Consistency over time. Two of the three cases demonstrate that influence depends not on a single dataset but on sustained measurement over time. Notably, the FRA has carefully balanced methodological innovation with comparability across successive iterations, and Mpact has adapted its methods to a rapidly changing technological and political environment while preserving continuity in community-led estimation.²⁹ These experiences suggest that consistency is not primarily a technical achievement but an institutional one. Data systems become influential when organizations develop the capacity,

partnerships, and governance arrangements necessary to sustain measurement beyond individual projects or funding cycles. Consistency over time also enables monitoring, allowing institutions and communities to track changes not only in LGBTI outcomes, but also in the participation and data agency that shape how evidence is produced and used.

From unofficial to official. The three governance models also illustrate different pathways through which data acquire institutional authority. Community demands became official national statistics through INEGI; the FRA's regional surveys have become authoritative evidence for European policymaking despite existing outside of national statistical systems; and Mpact's community-generated estimates were incorporated into UNAIDS Key Populations Atlas and PEPFAR planning documents. Together, the cases demonstrate that authority is not determined solely by who produces the data. Rather, authority is created through successful translation across institutions, combining methodological credibility, trusted partnerships, and sustained engagement with decision-makers.

Doing more with limited resources. Each initiative achieved influence despite resource constraints. INEGI maximized limited questionnaire space by prioritizing variables with the greatest analytical value. The FRA operates one of its smallest-budget surveys while producing one of its most widely used datasets. Mpact relied on freely available digital signals and community partnerships to develop low-cost methodologies capable of operating where conventional surveys were infeasible. These findings suggest that the principal constraint facing the LGBTI data ecosystem is not always funding itself, but the ability to strategically connect complementary institutional strengths. Partnerships can extend the value of limited resources.

Twofold impact. The three case studies demonstrate that impact extends beyond individual policy changes. ENDISEG institutionalized LGBTI measurement within Mexico's national statistical system; the FRA has informed successive European equality strategies and national statistical discussions; and Mpact has influenced HIV financing while strengthening community capacity to generate and use evidence. The cases therefore suggest that data agency can produce impact on two levels simultaneously: first, agency contributes to changes in policies, programs, and resource allocation; and second, agency strengthens long-term data availability through new institutional relationships, technical capacities, and enduring systems capable of producing future evidence.

Recommendations

Pulling together all the patterns across the four stages of the Data Value Chain, and considering the five cross-cutting findings from the case studies, numerous recommendations emerge to support data agency within the Data Value Chain.

For Donors:

- Fund LGBTI data initiatives, especially to LGBTI organizations and other stakeholders who have long-term vision to use the data to influence policy or other positive outcomes.
- Invest in funded research and evaluations that document data's policy influence over time.
- Support multi-year partnerships between national statistical offices and LGBTI organizations.
- Direct funding toward the translation gap, the uptake and impact stages of the Data Value Chain, especially where the resources are thinnest.
- Provide rapid-response or bridge funding for community organizations facing a sudden loss of data access, so that measurement capacity does not collapse with a single funding or policy shift.

For Civil Society Organizations (CSOs) and LGBTI Non-Governmental Organizations:

- Build internal data literacy and analytical capacity to move from data recipients to data co-producers and translators.
- Adopt shared metadata, definitions, and minimum interoperability standards where feasible to strengthen comparability across community-generated datasets.
- Build evidence-to-policy pathways into the design of data initiatives from the outset by identifying intended users, decision-making processes, and opportunities for policy uptake before collection begins.
- In some cases, sharing data with other organizations or researchers can allow those data to be further explored and used.
- Pursue partnerships with national statistical offices, regional organizations, and international institutions that create pathways for community-generated data to inform official systems and policy processes, while preserving appropriate community agency and ownership.

For National Statistical Offices:

- Where appropriate and safe, adopt tested and inclusive SOGIESC measures in relevant national surveys, drawing on existing guidance while allowing for contextual adaptation.

- Establish formal mechanisms for CSO and LGBTI NGO participation across the data value chain, not just at the collection stage.
- Implement privacy-by-design principles and safety protocols throughout data collection, storage, and access.
- Invest in secure, accessible microdata access mechanisms that allow researchers and community organizations, not only government users, to conduct their own analyses.
- Build context-appropriate data-integrity safeguards into politically sensitive LGBTI data collection.
- Use participatory prioritization to identify the most policy-relevant intersectional variables, balancing community priorities with analytical value, respondent burden, sample size, safety, and comparability over time.
- Design publication strategies for different users, including accessible community-facing products, policy-oriented outputs, and appropriately protected microdata, rather than relying on a single report or dataset.

For United Nations Specialized Agencies:

- Support community-led LGBTI data initiatives as part of broader inclusive data investments.
- Issue technical guidance on SOGIESC measurement for national statistical systems.
- Support cross-regional learning and harmonization that improves comparability while preserving the flexibility to adapt instruments to national context, community priorities, and safety considerations. In-person or virtual South-South Knowledge Exchanges offer one way to facilitate knowledge sharing.
- Develop and pilot approaches for monitoring LGBTI data agency, including changes in community participation, influence, and decision-making power across the Data Value Chain.

For Economic Development Institutions:

- Provide financial and technical assistance to national statistical offices undertaking data collection, including for LGBTI people.
- Support multi-year data partnerships between national statistical offices and CSOs and LGBTI NGOs.

- Collaborate with regional and international institutions to work as bridges between citizen data and national policy and statistical systems, particularly where direct pathways to national institutions are limited.
- Finance the various stages of the Data Value Chain directly, as well as support capacity building for evidence-to-policy pipelines, rather than funding data collection alone.
- Support contingency and data stewardship planning, including secure archiving and legal safeguards for LGBTI data systems operating in an increasingly hostile political environment, so that existing evidence and infrastructure are not lost to funding cuts or policy reversals.

For Academia and Research Institutions:

- Invest in methodological innovation for hard-to-reach LGBTI sub-group measurements.
- Develop intersectional analytical frameworks that go beyond bivariate disaggregation.
- Partner with community organizations as co-investigators throughout the research process, not solely as data sources.
- Document and publish negative or null cases where LGBTI data collection was attempted but did not translate into use or policy influence to help the field learn from the translation gap, not only from success stories.

Conclusion

The case studies demonstrate the future of intersectional LGBTI data depends not simply on producing more evidence, but on strengthening the relationships, institutions, and governance processes that enable evidence to influence policy. The analysis also shows that data agency can be a critical mechanism through which LGBTI data move through the Data Value Chain and generate public value. Data become influential not because they exist, but because communities, organizations, national statistical offices, researchers, and international organizations shape how evidence is collected, published, and ultimately used.

Viewed through this lens, strengthening LGBTI data requires more than better measurement. It requires sustained investment in partnerships, institutional capacity, governance, and the connections that translate evidence into policy, financing, and institutional change. It also requires recognizing intersectionality not only as greater disaggregation or the addition of more variables, but as a governance question about whose identities and experiences are represented, who participates in decisions about data, and what can be measured credibly, ethically, and safely.

The future of LGBTI data, however, cannot be taken for granted. In many parts of the world, political and institutional developments are challenging the legitimacy of collecting, publishing, and using data on LGBTI people. Statistical institutions, funding, and legal protections that support inclusive data systems have come under increasing pressure. Strengthening LGBTI data systems therefore depends not only on technical innovation, but also on protecting the legal and institutional policies that make safe, ethical, and participatory data systems possible.

Moving “beyond counting” ultimately means recognizing LGBTI communities not simply as subjects of measurement, but as crucially important participants in shaping how evidence is produced, governed, interpreted, and translated into more inclusive public policy and outcomes.

Endnotes

- ¹ ODW and Data2x, Harnessing the Power of Intersectionality Data for Better Development Policies and Impact: A Concept Note, (Washington, DC: Open Data Watch, 2025), <https://opendatawatch.com/publications/intersectionality-data-for-development-impact/>
- ² The process through which data are created and used, from identifying a need for data through collection, publication, uptake, and impact, including potential reuse. See ODW, The data value chain, <https://opendatawatch.com/reference/the-data-value-chain-executive-summary/>
- ³ The National Statistical Office or Institute is the government agency, office, or institution that serves as the core producer and primary national institution responsible for collecting, processing, and disseminating official statistics concerning the nation as a whole.
- ⁴ Elizabeth Lockwood, Intersectional Data and Persons with Disabilities (Washington, DC: Open Data Watch, 2025), <https://opendatawatch.com/publications/intersectional-data-and-persons-with-disabilities/>.
- ⁵ Clifton Cortez, Toni Joe Lebbos, and Lucas Regner, "Is Your Country LGBTI Inclusive? With Better Data, We'll Know," World Bank Blogs, December 11, 2017, <https://blogs.worldbank.org/en/opendata/your-country-lgbti-inclusive-better-data-we-ll-know>
- ⁶ Open Data Watch, Harnessing the power of intersectionality data for better development policies and impact, <https://opendatawatch.com/publications/intersectionality-data-for-development-impact/>
- ⁷ Ibid.
- ⁸ Ibid.
- ⁹ The Collaborative on Citizen Data's Copenhagen Framework defines citizen data as "data originating from initiatives where citizens either initiate or are sufficiently engaged, at the minimum, in the design and/or collection stages of the data value chain, irrespective of whether these data are integrated into official statistics." See, United Nations Statistics Division, The Copenhagen Framework on Citizen Data (New York: United Nations, 2025), <https://unstats.un.org/UNSDWebsite/citizenData/events/The-Copenhagen-Framework-on-Citizen-Data-new-design.pdf>
- ¹⁰ M. V. Lee Badgett, Intersectionality and SOGIESC Data: Opportunities and Challenges, Discussion Paper (Washington, DC: Data2X, April 2024), Data2X.
- ¹¹ The taxonomy was created by the authors and colleagues at Koppa Lab, publication forthcoming.
- ¹² Collaborative on Citizen Data, Guidance on Intersectionality (Collaborative on Citizen Data 2026), Chapter 4, Ethical Risks of Intersectional Data for Marginalized Groups, <https://unstats.un.org/UNSDWebsite/citizenData/resources/Guidance-on-Intersectionality-June2026.pdf>

- ¹³ Elizabeth Lockwood, Intersectional Data and Persons with Disabilities (Washington, DC: Open Data Watch, 2025), <https://opendatawatch.com/publications/intersectional-data-and-persons-with-disabilities/>.
- ¹⁴ The Mpact case study uses the term gay, bisexual, and other men who have sex with men (GBMSM), consistent with terminology commonly used in HIV programming and public health. We adopt that terminology when discussing the Mpact case study.
- ¹⁵ Mpact Global Action for Gay Men's Health and Rights, "Home," accessed July 7, 2026, <https://mpactglobal.org>
- ¹⁶ We adopt the EU-designated acronym "LGBTIQ" to specifically include queer-identifying people, when examining the FRA case study.
- ¹⁷ European Union Agency for Fundamental Rights, LGBTIQ Equality at a Crossroads: Progress and Challenges (Luxembourg: Publications Office of the European Union, 2024), <https://fra.europa.eu/en/publication/2024/lgbtiq-equality-crossroads-progress-and-challenges>.
- ¹⁸ Here we use LGBTIQ+ in line with INEGI's acronym, specifically including queer identifying people as well as others with diverse SOGIESC (+) who might not identify as LGBTIQ.
- ¹⁹ National Institute of Statistics and Geography (INEGI), National Survey on Sexual and Gender Diversity (ENDISEG) 2021, accessed July 7, 2026, <https://en.www.inegi.org.mx/programas/endiseg/2021/>
- ²⁰ Stefan Baral, Rachael M. Turner, Carrie E. Lyons, Sean Howell, Brian Honermann, Alex Garner, Robert Hess, Daouda Diouf, George Ayala, Patrick S. Sullivan, and Greg Millett, "Population Size Estimation of Gay and Bisexual Men and Other Men Who Have Sex With Men Using Social Media-Based Platforms," JMIR Public Health and Surveillance 4, no. 1 (2018): e15, <https://doi.org/10.2196/publichealth.9321>
- ²¹ European Union Agency for Fundamental Rights, LGBTIQ Equality at a Crossroads: Progress and Challenges (Luxembourg: Publications Office of the European Union, 2024), <https://fra.europa.eu/en/publication/2024/lgbtiq-equality-crossroads-progress-and-challenges>
- ²² European Union Agency for Fundamental Rights, Technical Report – EU LGBTIQ Survey III (Luxembourg: Publications Office of the European Union, 2025), <https://fra.europa.eu/en/publication/2025/technical-report-eu-lgbtiq-survey-iii-0>
- ²³ Here, "proximity" refers broadly to the degree of access, relationships and influence that data users have with decision-makers, rather than merely physical distance.
- ²⁴ ILGA-Europe, Intersections 2.0: Key Findings on LGBTIQ Health in Europe (Brussels: ILGA-Europe, February 4, 2026), <https://www.ilga-europe.org/report/intersections-2-0-key-findings-on-lgbtiq-health-in-europe/>

- ²⁵ amfAR, When Size Matters: How Social Media Can Help Determine Key Population Size (Washington, DC: amfAR, The Foundation for AIDS Research, July 18, 2018).
- ²⁶ Baral et al., "Population Size Estimation of Gay and Bisexual Men," e15.
- ²⁷ Ibid.
- ²⁸ Council of Europe, Recommendation CM/Rec(2025)7 of the Committee of Ministers to Member States on Equal Rights for Intersex Persons: Explanatory Memorandum (Strasbourg: Council of Europe, 2025).
- ²⁹ ENDISEG has been designed as a recurring survey, with upcoming plans for the second iteration.

Annex 1: Resources

The following resources are intended as practical guidance to strengthen LGBTI data agency throughout the DVC. Rather than general background reading, they offer methodological guidance, standards, and implementation tools by (1) civil society organizations, (2) academic institutions, (3) international organizations and networks, and (4) national statistics offices. Resources are organized first by the themes of data agency and intersectionality, then listed by the four stages of the DVC to help readers identify tools relevant at each stage.

Foundations for Understanding LGBTI Data Agency

While the concept of data agency has emerged within broader discussions of inclusive data, it has not yet been systematically applied to LGBTI data systems. A broader examination of data agency can therefore be useful.

- From civil society, Open Data Watch and Data2X offer a framework for integrating intersectional data into official data systems and across the DVC, in [Harnessing the Power of Intersectionality Data for Better Development Policies and Impact](#). This Concept Note introduces data agency as a core principle and argues that communities (especially marginalized groups) should actively shape, govern, and use data that reflect their lives to influence policy.
- Also from civil society, Liga de Salud Trans in Colombia frames “statistical justice” as the right to be recognized within state systems and for official data to reflect the true diversity of society. In [El derecho al futuro: la justicia estadística como condición para proteger las vidas trans y no binarias](#), Liga argues that data should help identify inequalities, allocate budgets, design public policies, evaluate state action, and enable accountability, while ensuring that data do not become a weapon against transgender and non-binary people. See more on its website: <https://statisticaljustice.com>.
- From academia, a special issue of Big Data & Society offers five perspectives on the centrality of agency within data processes. [The introduction](#) to the special issue on data and agency in the Big Data & Society journal summarizes those perspectives “as data acquire new power, it is vital to explore the space for citizen agency in relation to data structures and to examine the practices of data work, as well as the people involved in these practices.”
- From international networks, the Global Partnership for Sustainable Development Data published a paper on data agency, accountability, and action. In [Reimagining Data and Power: A roadmap for putting values at](#)

[the heart of data](#), agency is defined as “having power to shape personal and/or community data and deciding whether, when, and with whom to share it”.

For a programmatic approach that focuses on LGBTI data agency, Koppa Lab launched Data LIFT (Data on LGBTI Inclusion for Transformation) in [Belize](#), as well as [Colombia](#), Mexico, and Brazil. Data LIFT is an action-oriented framework to translate data on LGBTI people into policy and financing pathways. Six steps guide the program in each country: (1) partner and create a coalition to cultivate agency among civil society organizations to work with data, (2) pull together existing data from various sources, and analyze it, (3) present findings in Data Roundtables, (4) work with the coalition to use findings to influence policies, (5) target international organizations and the development banks and use findings in their processes to unlock finance, and (6) partner with NSOs to establish larger data collection efforts in the future.

On Intersectionality and Data

With strong roots in Kimberle Crenshaw’s paper “[Mapping the Margins: Intersectionality, Identity Politics, and Violence Against Women of Color](#)” (1991), the concept of intersectionality refers to multiple identities that can overlap to shape the experiences of individuals or groups of people. This starting point is foundational to the DVC, and an application of an intersectional lens is important to understand data collection for LGBTI people.

- From civil society, Data2X published a discussion paper on the opportunities and challenges of integrating SOGIESC into intersectional data systems. The paper “[Intersectionality and SOGIESC Data: Opportunities and Challenges](#)” also provides practical guidance for collecting and analyzing data on LGBTI people.
- From international networks, the Collaborative on Citizen Data published “[Guidance on Intersectionality](#)” which demonstrates how intersectional data can be translated into more equitable policies by helping governments better understand and address the experiences of people in all their diversity.

Data Collection

As the foundation of the DVC, the collection stage determines what data are collected and how and sets the trajectory for subsequent stages. As the most mature stage of the LGBTI DVC, it has generated a substantial body of guidance from many stakeholders.

- From civil society, in 2022 the National Academies of Sciences, Engineering, and Medicine launched [Measuring Sex, Gender Identity, and Sexual Orientation](#) to give recommendations for collecting sex and SOGI data, especially to guide more consistent and inclusive measurements across statistical systems.
- From academia, and specifically regarding gender identity, the Williams Institute presents best practices for collecting data on transgender and other gender minority people, including the widely used two-step approach to asking sex assigned at birth and current gender identity, in [Best Practices for Asking Questions to Identify Transgender and Other Gender Minority Respondents on Population-Based Surveys](#).
- From international organizations, the EU's FRA (drawing from the case study) published its [Technical Report – EU LGBTIQ Survey III](#) as well as the questionnaire, which can guide the creation of online, large surveys covering many countries.
- In 2019 the United Nations Independent Expert on Sexual Orientation and Gender Identity published [Data collection and management as a means to create heightened awareness of violence and discrimination based on sexual orientation and gender identity](#) to advance a human rights framework for collecting and managing SOGI data, especially emphasizing ethical practices.
- From national statistics offices, both Australia and New Zealand created standards of SOGIESC measures to incorporate into state-led data collection efforts. In 2020, the Australian Bureau of Statistics published the [Standard for Sex, Gender, Variations of Sex Characteristics and Sexual Orientation Variables](#) and in 2021, Stats New Zealand published the [Data Standard for Gender, Sex, and Variations of Sex Characteristics](#).

Publication

The following resources focus on publication, where collected data become visible through reports, datasets, indicators, and other knowledge products.

- From civil society, the [Global Philanthropy Project](#) and Funders for LGBTQ Issues jointly publish the Global Resources Report, the most comprehensive dataset on government and philanthropic funding for LGBTI people around the world. The report aggregates grant records into a structured, comparable knowledge product for advocates and funders to use in their work. The 2023–2024 edition tracks over 19,000 grants from more than 1,300 foundations and government agencies. Although this report does not focus on data collection itself, it is a robust example of accessible data.

- From academia, INED and Leiden University published the [LawsAndFamilies Database](#), multiple knowledge products showing the legal, statistical, and sociological perspectives on same-sex marriage and relationships in the European region. Two of those knowledge products include a legal database tracking 60 legal aspects of marriage, partnership, and parenting across 21 countries since 1965, and a compilation of statistical data on same-sex marriage in 12 countries.
- From international organizations, UNDP suggests a standardized set of indicators for measuring LGBTI inclusion across five domains. In [A Set of Proposed Indicators for the LGBTI Inclusion Index](#), UNDP provides a framework for the publication of consistent, comparable data that can be used to monitor progress and guide evidenced-based policies.
- Also from international organizations, the World Bank provides a standardized framework of legal and policy-based indicators for benchmarking LGBTI inclusion. In [Equality of Opportunity for Sexual and Gender Minorities](#), or EQOSOGI, the report provides a comparative evidence base to support monitoring and policy dialogue across countries.
- From national statistical offices, (drawing from the case study) Mexico's INEGI and [its ENDISEG website](#) produce a broad array of publicly accessible materials, including the survey report, methodology, questionnaire, metadata, statistical tables, and dissemination products.

Uptake

The following resources can guide on the uptake stage, or how data are interpreted, shared, and translated into decision-making processes.

- From civil society (and pulling from the case study), Mpact wrote a technical brief to guide HIV advocacy groups to understand and engage with the Global AIDS Monitoring (GAM) framework. In "An Introduction to Global AIDS Monitoring for Key Populations," Mpact shows how to use official HIV data for advocacy, accountability, and better decision-making. See more here: <https://www.mpactglobal.org/publications>.
- From academia, [Queer Data: Using Gender, Sex and Sexuality Data for Action](#) offers insights on select data collection initiatives in the U.K. and suggests how they can inform policy, services, and advocacy.

Impact

The impact stage refers to the ultimate goal of the Data Value Chain, where data are translated into tangible improvements in policies or other beneficial outcomes. Compared with the previous stages of the DVC, relatively few

resources were found that document the downstream impact of LGBTI data on policy or other interventions.

- From civil society, the Movement Advancement Project (MAP), publishes “[LGBTQ Equality Maps](#)” in the US, that translate legal and demographic data into a tool to help policymakers, advocates and others monitor policy change, gaps, and legislation. These visualizations show how laws and policies directly impact the lived experiences of LGBTQ people.
- From international organizations, the Council of Europe's [Recommendation CM/Rec\(2025\)7 on Equal Rights for Intersex Persons](#) shows how disaggregated survey data (including the FRA's LGBTIQ survey data) translated into the first comprehensive international legal instrument on the rights of intersex people, providing guidance to the Council of Europe's 46 member states.

Annex 2: Interview Questionnaire

Opening and Informed Consent

Opening script

Thank you for agreeing to this interview. As background: this report is commissioned by Open Data Watch (ODW), and is a collaboration between ODW and Koppa Lab. The report explores how data agency for LGBTI people and communities is exercised across the Data Value Chain. The conversation will take about 60 minutes, and, with your permission, we would like to record it for transcription and analysis, only accessible to us [Philip Crehan and Flavio Carrera Verdesoto]. Any quotes will be sent back for your validation prior to publication. We will also follow up with the final publication and plans for targeted dissemination.

Do you have any questions before we start?

Do you consent to record this interview and to proceed?

Background questions

- To begin, can you briefly introduce yourself, your role, your organization, and your organization's role in producing or using LGBTI data?
- Tell us about the data initiative we will discuss today: what is the work, when did it take place (or is it ongoing), where, and how many people were involved?
- From your perspective, what gap or problem did this initiative aim to address, and why was it important to do so through data?

Data Collection

Common questions

- What was the methodology of data collection (instruments, sampling, modality), and who led the key design decisions? Where were the disagreements and how were they resolved?
- How were priorities, gaps, and the characteristics and intersections of the data determined? Specifically, how were the SOGIESC variables defined, and what intersecting identity dimensions did you capture (for example race/ethnicity, age, disability, migration status, geography, class)?
- Did the community co-create or co-design data operations with key partners, such as national statistical offices? At what stages did LGBTI people and organizations have real decision-making power, and where was their role only consultative or absent?
- What safety, confidentiality, informed consent, and duty-of-care protocols did you put in place — for respondents and for field staff? Were there any incidents, and how were they handled?
- What barriers (legal, technical, institutional, financial) did you encounter, and how did you navigate them? How was the data collection funded, and did the funding source shape what you could ask or how?

Adaptation — INEGI:

- What internal and external dynamics enabled INEGI to commit to a dedicated LGBTI instrument (ENDISEG and its iterations)? What roles did CONAPRED, national LGBTI organizations, and academia play in reaching that decision?
- How was the SOGIESC module developed: what informed the wording and response options, what was included, and what was left out? Were there cognitive pretests with LGBTI respondents?
- What formal or informal mechanism for LGBTI civil society participation exists in INEGI's processes, and how binding is it?

Adaptation — Mpact:

- In contexts where same-sex relations are criminalized or LGBTI organizing is restricted, how do partner organizations on-the-ground design and run data collection safely?

- What role does Mpact play vis-à-vis the community-led organizations doing the direct collection? How do you protect their agency rather than substitute it, and where are the real tensions between technical support and capture of the agenda?
- How is HIV-related data (prevalence, cascade, service access) integrated with broader LGBTI experiences (violence, mental health, criminalization, displacement)? What is measured together, what is kept separate, and why?

Adaptation — FRA:

- How do you balance methodological standardization across multiple countries with national or community specificity? What is held constant across rounds (2012, 2019, 2023) and what is adjusted?
- How are LGBTI organizations across member states involved in the design? When disagreements arise — between national organizations, or between them and FRA’s research team — how are they resolved?
- Which intersecting dimensions does the survey capture, and which were left out? What concrete trade-offs did you face (between race/ethnicity, migration status, or disability, for example)?

Data Publication

Common questions

- What was published, in what formats, and at what level of disaggregation? Who decided what to release and what to hold back? Were there findings you chose not to publish, and why?
- How are intersectional findings presented when sample sizes for crosscuts are small or re-identification risk is high?
- Who has access to the underlying microdata, and under what conditions? Do LGBTI communities or organizations have the same level of access as other research users?
- What political, technical, or legal challenges have you faced in publishing LGBTI data? Have any findings been suppressed, delayed, or modified?
- How do you communicate findings back to LGBTI communities themselves? And in what formats, languages, and platforms?

Adaptation — INEGI:

- How does INEGI handle sensitive findings of bias-motivated violence, mental health, experiences with state institutions, in its public products? Are there differences between the press release, the tabulated outputs, and the microdata?
- What is INEGI's microdata access policy for LGBTI researchers and CSOs? Is there friction between confidentiality protocols and community demand for own analysis?

Adaptation — Mpact:

- When partner organizations operate in hostile legal environments, how is it decided what Mpact publishes globally, what partners publish locally, and what stays internal? How is attribution handled to protect contributors?
- What happens when community data contradicts official figures (governments, WHO/UNAIDS) on prevalence or coverage? How is that discrepancy published, and who bears the reputational cost?

Adaptation — FRA:

- How is it decided what enters the headline report vs. the data dashboards vs. the country fact sheets? What rules govern reporting at member-state level when sample sizes are small or politically sensitive?
- How are findings communicated to national audiences (governments and CSOs), and what feedback do you receive on the usefulness of these products?

Data Use

Common questions

- Who has used these data for policy, programming, advocacy, litigation, research, resource allocation, or other purposes?
- Can you describe a concrete case in which the data you produced was taken up by a decision-maker (government, donor, multilateral, court)? What made that uptake possible?
- How have your own organization and the LGBTI communities involved used the data? Is there a notable difference between what the community can do with the data and what third parties (governments, donors) can do with it?

- Can you give an example where findings led to a different conclusion or response than aggregate LGBTI data would have?
- Where has the data not been used as you hoped, and what blocked that use? What data do you wish existed that currently does not?

Adaptation — INEGI:

- How have ENDISEG and successor instruments been used by CONAPRED, sectoral ministries, sub-national governments, and civil society? Is there a feedback loop where users shape future rounds?
- Is there evidence of the data informing budgets, programs with LGBTI registries, or legal reforms? What about individual cases (strategic litigation, CNDH recommendations)?

Adaptation — Mpact:

- How does community data feed into PEPFAR, the Global Fund, UNAIDS, and bilateral programming? At what points in the funding cycle (strategic, allocation, monitoring) does it gain traction, and where does it not?

Adaptation — FRA:

- How has the EU LGBTIQ Survey been used by the Commission, the Parliament, member states, and European CSOs (for example, ILGA-Europe)? Are there legal or policy changes traceable to specific findings (such as the LGBTIQ Equality Strategy 2020–2025)?
- How is the data used for infringement procedures, rule-of-law monitoring, or pressure on member states experiencing backsliding (for example, Hungary, Poland)?

Impact

Common questions

- How do you define and measure impact in this kind of data work? Who decides what counts as impact, your organization, funders, or the community?
- Can you point to a tangible change (in policy, law, programs, public discourse, or resource allocation) that this data contributed to? What evidence do you have for that attribution?
- What enabling conditions (political, institutional, technical, financial) made that impact possible? What blocked it where it did not occur?

- Beyond the published outputs: did the process of producing the data itself shift anything, for the LGBTI community involved, for the host institution, or for the relationship between them?
- With hindsight, what would you do differently if you started this initiative today, particularly around community leadership and ownership of the process?

Adaptation — INEGI:

- Has the existence of ENDISEG changed how other Mexican institutions (federal, state, judicial) collect or use SOGIESC data? Has it had demonstrative effects in the LAC region?
- What changed inside INEGI as a result of doing this work, in capacities, in institutional culture, in relationships with civil society?

Adaptation — Mpact:

- How has community-led data shifted the way global HIV institutions engage with GBMSM populations, in funding, programming, operational definitions, and governance mechanisms?
- What effect has it had on the in-country partner organizations themselves? For example, on their negotiating power vis-à-vis governments and donors, on their sustainability, and on their safety?

Adaptation — FRA:

- What is the documented track record of EU institutions and member states acting on FRA findings; including instances where they did not act, and what that gap suggests about the limits of evidence?
- How has the LGBTI data ecosystem in Europe changed because of the survey's recurrence, in national surveys, academic research, and CSOs?

Closing

- Looking across the entire Data Value Chain in your work: where would you say LGBTI agency is strongest, and where is it weakest? Why?
- Is there anyone else we should be speaking with for this report — a person, organization, or case that illuminates something we are missing?
- Is there anything we did not ask that is important for understanding this work?