Guide to Equity-Driven Data Management

Tools for Developing a Gender-Based Analysis Plus-Enabled Data Management Strategy

AUGUST 2024

AMSSA®
Strengthening Diversity In BC

Contents

Contents

Acknowledgments

Addressing systemic inequities and advancing gender equality for all newcomers is central to their successful integration.

Immigration, Refugees
and Citizenship Canada

Immigration, Réfugiés
et Citoyenneté Canada

Introduction

Through *The Enhancing National Sector Gender-Based Analysis (GBA) Plus Capacity* project, umbrella organizations across Canada have partnered to enhance the immigrant and refugee serving sector's capacity to mainstream GBA Plus within programs, policies, and procedures.

Data is an essential component of any GBA Plus initiative, as it enables the identification of differences in how clients access and benefit from services. Newcomer-serving agencies often possess large amounts of data, but do not always derive as much value from that data as is needed or wanted. This document is designed to guide newcomer-serving organizations through the process of developing and implementing a GBA Plus equity-driven data management strategy. Each step is laid out and accompanied by supporting worksheets.

A comprehensive data management strategy allows organizations to ensure their data practices align with organizational mission, vision, and values; enhances equity work; and empowers agencies to derive maximum value from their data throughout its lifecycle. This guide can be used to inform organizational data management no matter where the organization is at in their data management journey. However, regardless of the stage an organization is at, it is recommended that the agency take the time to identify their data needs and goals, as outlined in Phase One.

Any organization serving newcomers can use this guide. Although it provides specific steps and questions for each stage of the process, it is intended to support internal processes, not to supersede them. GBA Plus and data equity depend on responsiveness to the context in which they are developed; therefore, organizations are encouraged to adapt the suggestions and guidelines contained here to fit their own work.

Because each agency engages with data in different ways and comes to the work with different knowledge and tools, this guide begins with an introduction to GBA Plus and data equity as well as an explanation of key data terms. There are many facets to data management, and while it is recommended that all data practices be examined, resource limitations may require agencies to prioritize certain areas for more in-depth evaluation. The following sections will help organizations make strategic decisions about where to focus their efforts.

Introduction

How to use this guide

This guide walks organizations through the foundations of data management and five subsequent phases of developing a data management strategy. Each section includes worksheets to support that phase as well as case studies illustrating their use. The chart on the following page provides a visual reference for the process of developing a data management strategy that fits organization needs, upholds data equity and organizational values, and enables GBA Plus analysis.

Foundations: In this section you will find a brief introduction to the importance of data for service-providing organizations (SPOs), information about GBA Plus in the newcomer-serving sector, six principles of data equity, and a glossary of terms for data work. These concepts are the foundation for the rest of this guide.

Phase One is identifying how the agency uses data through key questions, outlined on Worksheet 1. Because data is a tool for conveying meaning, this phase also asks what stories the organization wants to tell with their data and what data is needed to communicate each narrative to its respective audience; Worksheet 2 provides a template for this process.

Phase Two guides agencies in evaluating their current data practices to examine where they already uphold GBA Plus and data equity, see where practices meet the agency's data needs, and identify gaps and incongruencies. Worksheet 3 provides reflection prompts on equity within data practices, while worksheets in Appendix B guide evaluation of individual practices.

Phase Three supports agencies in using their data evaluation and identified data needs and goals to create a data management strategy that builds on existing strengths to fulfill needs and obligations while maintaining equity commitments. Use Worksheet 4 to design practices for meeting data needs and Worksheet 5 to integrate GBA Plus, values, and data equity.

Phase Four offers suggestions and considerations for agencies as they put their data management strategy into practice. It includes logistical questions as well as ones highlighting equity during the implementation process. Through Worksheet 6 you will establish an implementation plan for the overall strategy, while Worksheet 7 frames individual data practices.

Phase Five provides guidance for establishing mechanisms to regularly check data practices for equity, utility, and compliance with standards set during the Design phase. Worksheet 8 offers framing questions for monitoring data equity, while Worksheet 9 focuses more on logistics.

Finally, the **Conclusion** brings the answers reached through the various worksheets to a sample template and draft for a comprehensive data management strategy. Worksheet 10 is a complete template for agencies to adapt to their own needs when compiling a data management strategy based on their use of worksheets in Phases 1 through 5.

Introduction

How to use this guide

Appendix A consists of a more in-depth guide to the core data practices and related activities outlined in **Phase Two**.

Appendix B is a set of worksheets building on **Phase Two** by providing guiding questions for evaluating each core data practice. These worksheets support the potential activities described in Appendix A and provide useful information for designing a data strategy.

Phases of Building a Data Management Strategy
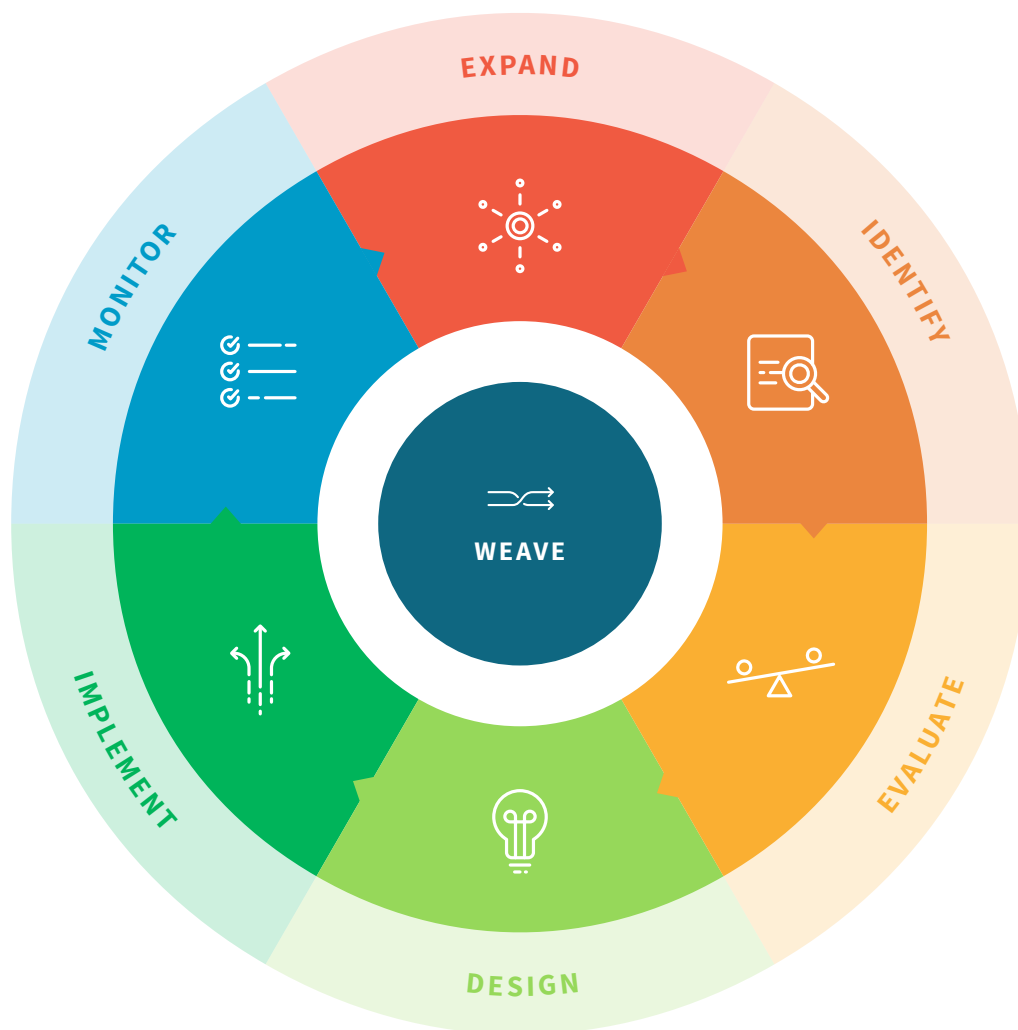
Data matters

All agencies use data all the time. Data plays a key role in many core functions within an organization.

It is used to describe the work being done, such as through reports to funders or communities; to evaluate programs and pinpoint successes and gaps; and to provide supporting evidence that demonstrates and proves what is seen and known internally. When front-line staff see that programs are not aligning with client needs, or when clients express barriers or challenges, data provides essential evidence to support efforts to improve, adjust, or expand programs and policies.

For example, staff may notice a drop in language class attendance during the summer months. Data helps us to understand precisely what sorts of clients are affected. Recording attendance, enrollment, and withdrawal numbers allows the agency to prove that what they observe is reflected in actual attendance numbers. Access to demographic data like age, gender, immigration status, family composition, and so on, allows the agency to pinpoint who is withdrawing from classes at this time. Examining that client data may uncover the fact that most of those withdrawals are parents with younger children. Now, instead of broad measures like distributing a survey to all language class students, the agency can focus their investigation on parenting clients. Engaging directly with withdrawing students leads to a greater understanding of the barriers they experience; for example, they may be unable to attend because they need to be at home taking care of their school-age children and cannot afford childcare that would allow them to attend classes. With this information, the agency is able to explore and design program interventions with the potential to remove those barriers – perhaps by connecting clients to affordable childcare programs, offering on-site childcare, or creating an in-home or remote tutor program for the summer months. Data like this can also be important when applying for funding, as it allows the agency to demonstrate the need they are seeking to meet (for example, funding that would allow them to offer these interventions). This sort of focused intervention is much more difficult to achieve without adequate access to data that helps explain and direct deeper investigations.

> **USING THIS GUIDE**
>
> At the beginning of each section, you will find a blue box just like this one with a brief overview of what you will find in that chapter and suggestions about how to get the most out of it.
>
> **In this section, Foundations, you will find:**
> 1. An overview of data work within GBA Plus, including data equity.
> 2. Key terms that will be used throughout the guide.
> 3. An introduction to the two hypothetical organizations whose data process will be presented as case studies in each phase.
>
> You can read these sections closely before embarking on the full process, or you can skim through this chapter and revisit its contents as you work through the worksheets and case studies in the rest of the guide. Both are valid; choose the approach that works best for you.

To get the most value out of data, it is essential to have a data management strategy based on how the organization uses – and hopes to use – their data. Because data can serve many purposes and be used in many different ways, a formal strategy begins with identifying needs and goals, a process outlined in Phase 1. But it can be dangerous to look at data solely as a tool or resource. Because the data that is most useful represents actual people, it is necessary to consider how data practices not only support organizational needs but also how clients are impacted by the ways data is collected and used. This guide uses GBA Plus and data equity as complementary frameworks to ground each stage of data management in equity, justice, and inclusivity.

GBA Plus and data equity

COMMUNICATE	**IDENTIFY ISSUE**	**DOCUMENT**
	CHALLENGE ASSUMPTIONS	
	GATHER THE FACTS—RESEARCH AND CONSULT	
	DEVELOP OPTIONS & MAKE RECOMMENDATIONS	
	MONITOR & EVALUATE	

EXCELLENT RESULTS FOR DIVERSE CANADIANS

Graphic drawn from Women and Gender Equality Canada's Introduction to GBA Plus online course

This guide and the accompanying worksheets provide agencies with a framework for evaluating and improving their existing data management practices in order to facilitate the development of a data management strategy that meets their needs and obligations. This process is guided by two lenses: Gender-based Analysis Plus (GBA Plus) and data equity.

GBA Plus is an analytical framework used to examine policies, programs, and practices to understand how different groups experience and benefit from policies and programs in different ways. The Government of Canada calls for its use in all policy and program design and evaluation, but engaging in GBA Plus analysis requires well-organized and robust data because data is the first and most foundational step in the process. Women and Gender Equality (WAGE) Canada describes GBA Plus as a process in which the first three steps are identifying issues, challenging assumptions, and gathering relevant facts.

Foundations

Data is essential to these three activities, as accurate understanding of people and situations is necessary to analyze how different groups use services and how well services meet their needs.

This requires data not just on service usage, but also the people using services, to be able to understand how factors like ethnic and linguistic background, racialization, gender identity, dis/ability, sexual orientation, among many others, interact to shape their experiences and needs. To carry forward the example about LINC student withdrawals, GBA Plus analysis would enable the agency to explore whether factors like racialization and disability play a role in students' ability to attend classes, further refining the understanding that would allow them to, for instance, offer tailored programs facilitating access for disabled students. Data is also helpful for challenging assumptions and checking bias, as it can demonstrate where those assumptions fail to reflect reality. The evaluation framework outlined in Phase 2, and expanded upon in Appendix A, helps agencies determine where their data practices support GBA Plus and where work may be needed to expand this capability.

Data equity is a complementary framework to guide anyone who works with data through a lens of justice, equity, and inclusivity. It builds on concepts of data *ethics*, which focus on appropriate use of data through privacy protections, recognizing and mitigating biased data, and ensuring data quality to avoid spreading misinformation.[1] Data *equity* considers how data is collected, analyzed, interpreted, and distributed through questions about power, bias, and discrimination. It acknowledges and seeks to minimize biases and the exclusion of underrepresented groups at all stages of the data process – collection, curation, analysis, and use.[2] Because data plays such a key role in program and policy development, including funding, data practices have an unavoidable impact on how individuals and communities access and benefit from services. Equitable data processes contribute to equity in resource distribution and outcomes for diverse groups. This guide includes worksheets to support agencies in using a data equity lens when examining their data practices.

Principles of data equity

Although data management can be a complex process, data equity is rooted in six foundational principles that can be used to ground and guide the development of a data management strategy.

Use the following principles to shape an equity-driven foundation for your agency's data policies and practices, along with these worksheets:

Worksheet 3 | Worksheet 5 | Worksheet 6 | Worksheet 8

1

PRINCIPLE 1:
DATA REPRESENTS PEOPLE.

Respecting people means respecting their data.

The decision to collect a piece of information makes a claim about the importance of that data. For example, to collect gender identities is to say that gender identity is important information to have, but the choice of what labels to offer also makes a statement about which identities are institutionally recognized and therefore valid. For this reason, it is important to consider the labels and categories with which various communities self-identify. Furthermore, treating information as valuable means putting it to carefully considered use. Data must not be manipulated in ways that may misrepresent the human beings that it describes, and must be used in a manner that is respectful to the people who have shared it.

2

PRINCIPLE 2:
DATA IS SUBJECT TO BIAS.

Check the data and let the data check you.

All humans experience conscious and unconscious bias, and processes can carry those biases even in the absence of intention. Bias can be introduced at any point in the data lifecycle and must be consistently guarded against. The decision of what questions to ask and how to phrase those questions can be affected by bias, through word choice and the decision to include or exclude certain categories. The act of collecting data can also be subject to bias and assumptions, as collectors may explain questions and categories in ways that affect client responses or even fill in answers without client input. Quality management in the curation process may also be affected by bias, such as the choice to exclude all data records missing specific pieces of data, or to lump categories together when response numbers are small. Even if great care is taken throughout all of these steps, analysis is subject to bias when analysts approach the data with preconceptions or fail to account for broader contexts, such as historic disadvantages faced by a specific community. GBA Plus analysis is hampered when data does not adequately represent the diverse groups involved. One way to counter bias is diversity among staff who interact with all stages of the data lifecycle, ensuring that they can safely express concerns and opinions which may contradict a more dominant view. Another is to compare quantitative and statistical data to the lived experiences of clients.

3

**DATA HAS CONSEQUENCES -
"WHAT GETS COUNTED COUNTS."**[3]

*Always ask what your
data is doing.*

When data accurately and fairly represents people, it contributes to equitable distribution of resources, and to policies and programs that are able to serve and empower diverse groups. When data misrepresents a population, that group may be harmed, through the perpetuation of stereotypes, reduced access to and lesser benefits from programs intended to serve them, the diversion of resources away from their needs, and silencing as their lived experience is overwritten by data-derived narratives. It is also possible for visibility to be harmful, exposing individuals to discrimination. Both sides of visibility must be considered. Newcomer-serving agencies provide invaluable support to immigrants and refugees, many of whom exist at the intersection of multiple axes of marginalization. It is extremely difficult to design programs and policies that serve vulnerable groups if the information used for program and policy development is inaccurate or misused, and it is essential to listen to clients and communities to learn how data practices are impacting them.

4

**PRINCIPLE 4:
DATA DOES NOT EXIST IN A VACUUM.**

*Context shapes understanding
and interpretation.*

People exist at a confluence of place and time, and their experiences are shaped by both past history and current conditions. Many communities face or have historically faced discrimination, oppression, and other disadvantages. Therefore, it is essential to place data in its proper context. For example, racialized newcomers may have lower employment rates compared to non-racialized newcomers. This piece of data must be considered in the context of racial disparities in access to education and training, discriminatory or biased hiring practices, and available opportunities in the country of origin. If use of employment services is also lower among some of these groups, that must also be examined in context. Lived experience is valuable when contextualizing data, because data such as service use statistics can demonstrate where a disparity exists but cannot, on its own, explain why that disparity exists or how it could be addressed.

5

**PRINCIPLE 5:
DATA MANAGEMENT MUST BALANCE
PUBLIC GOOD AND PERSONAL PRIVACY.**

*Be intentional about what you collect
and how you protect it.*

All data collection is invasive, as it asks individuals to disclose information they may not ordinarily share. Some data may be extremely sensitive, such as for LGBTQIA+ persons who are not "out." Every piece of data must be collected for a purpose, with informed consent, in a culturally safe manner, and used to benefit the public good. All data must also be carefully protected to ensure privacy and confidentiality, including anonymization where feasible.

6

**PRINCIPLE 6:
DATA MANAGEMENT IS A LEARNING PROCESS.**

*Commit to owning and learning from
mistakes made along the way.*

Data practices are designed and carried out by human beings, with all the imperfections of any human endeavor. At times, these practices will cause harm or introduce errors. For example, an error in collection practices may result in policies that overlook the access needs of a specific community, or in data analysis algorithms that reproduce harmful stereotypes. In such cases, an agency must be ready to acknowledge the impacts of this error, assess their practices to determine its cause, and remediate those practices so that they are more equitable in the future.

Key Data Terms

AGGREGATION is the process by which information about multiple people, such as the entire client base of a program, is compiled from databases to prepare combined data sets for processing. Aggregation is necessary to analyze groups of people. *See also:* Disaggregation.

COLLATION brings together different pieces of data for comparison. This is often done after disaggregation, so that similarities and differences between groups can be identified.

CULTURAL SAFETY is the creation of a space where clients feel safe disclosing their personal information and is necessary for informed consent and collecting accurate data.

CURATION is the term for processes used in processing, housing, protecting, and maintaining data in order for value to be derived from it.

DATA EQUITY describes "the principle of fair and just access, representation, and use of data across diverse communities and stakeholders."[4]

DATA MANAGEMENT describes how an agency collects, stores, curates, and uses data. A data management strategy is an important tool for ensuring that all steps in the data process align with the agency's mission, vision, and values as well as answer key questions with quality data.

DATA NARRATIVES are a framing device for organizing and analyzing data around a specific purpose, using evidence tailored to speak to a particular audience.

DATA QUALITY refers to the accuracy, completeness, and consistency of data. Inaccurate data may include invalid data points, missing values, or outlier data.

DATA STANDARDS are the rules used to standardize how data is described, represented, and structured, such as common terminology, consistent representation of dates, and shared definitions. It also includes making data recognizable through an accessible database or file format. Data standards ensure that data is findable, accessible, easy to use/re-use, and share.

DATA SETS are collections of data. Often, an individual dataset is created by pulling specific pieces of data from a larger collection, such as Amharic-speaking refugees from an agency's data on all clients, or temporary foreign workers accessing a specific program. **Predetermined datasets** are created by programming a database to return a specific subset of data in a reusable manner.

DE-IDENTIFIED data is created by the removal of personal identifiers, such as name, SIN, personal health number, and address. This protects the privacy of the people from whom data is collected and is sometimes referred to as **anonymizing** data, or making it anonymous. It is different from **confidentiality**, where personal identifying information is protected but remains linked to all of a person's data.

DISAGGREGATION is the process by which aggregated data is separated by factors such as gender, race, immigration status, or dis/ability, among many others. It is essential for GBA Plus analysis and equity-driven work because it enables identification of differences in how various client groups access and benefit from services. *See also:* Single-factor and Multi-factor analysis.

Key Data Terms

INTEROPERABILITY describes the ability to share data between stakeholders or databases and the ease with which it can be combined with other data. For data to be interoperable, they must share equivalent standards (such as age ranges) and use compatible file formats (e.g. CSV).

LONGITUDINAL DATA describes the manner by which iterations of data are linked to create a picture of change over time. This form of analysis is most useful for tracking outcomes, trends, and impacts. *See also:* Snapshot data.

METADATA is information provided alongside data to make it easier to use, such as the source and how it was cleaned, processed, and/or calculated. It may also include a description of the data quality, such as missing data, how the data was validated, and what data standards were applied or followed. It can be thought of as a **data biography**, describing how the data came to exist in its current form.

MULTI-FACTOR ANALYSIS is a method for enabling **intersectional** analysis through comparison. An example is looking at outcomes for refugee women (intersecting factors: gender and immigration status) and comparing that information to outcomes for other groups to see how gender and immigration status interact to affect the experiences of people situated at that identity intersection.

QUALITATIVE DATA is information derived from non-numeric sources, such as client testimonials, interviews, or surveys with open-ended questions. It can provide context to **quantitative data**, which uses statistical analysis to answer questions. In particular, quantitative data is helpful for identifying *where* disparities are, but qualitative data is where investigators learn *how* and *why* those disparities exist.

SINGLE-FACTOR ANALYSIS examines how one identity factor influences client experiences. For example, analysis by gender is single-factor. This is helpful to determine disparities, though it is important to also consider the possible effects of other factors, such as race or immigration status, through multi-factor analysis.

SNAPSHOT DATA provides a picture at a single point in time. This might be cumulative (such as a total number of clients who have accessed services to-date) or iterative (such as how many active client files exist at a specific point in time). This sort of rapid data picture is best generated through pre-determined datasets.

WAREHOUSING is the manner in which data is stored. This describes both the location (database or client management system, server, cloud server, etc.), the protections that are in place to protect privacy, and the means by which the data is accessed, including who has permission to view, change, or export data.
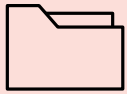
Case Studies

Because data work is complex, it often helps to have access to examples of how it could be done.

Integrated into this guide are two case studies providing glimpses into how this guide and its worksheets can be used in different contexts. Look for the **RED** and **YELLOW** color-coded boxes to see how two hypothetical organizations navigate the data management strategy development process, from identifying their data needs all the way through monitoring their ongoing work. Use these case studies to get a sense for how the work *could* be done, understanding that some pieces may be more relevant to your organization's context than others. The next few paragraphs offer an introduction to the context for each case study. Note that these case studies use fictional organizations; examples are composites and extrapolations loosely drawn from multiple organizations who tested this guide.

CASE STUDY A

ORGANIZATION A is a small settlement service provider serving several small communities in a rural area.

12 staff	Regional, rural service area	Recently digitized	Starting to build data practices

They made the shift from paper files to a digital database less than a year before embarking on a data management strategy design. Currently, their data collection is narrowly focused on reporting to funders, but as they have worked to integrate GBA Plus into their practices, they find that they need access to more data to dig as deeply into their work as they would like. Although they feel confident in their current knowledge about their work, they also have a number of questions that they feel data can answer, and they want to be able to support and challenge their rich qualitative data with quantitative data.

Equity is a driving force within this organization, and they are very intentional about how they engage with their work, both when working with clients and when working with each other. As with any organization, their priorities and concerns are shaped by the context in which they work. In this case, they connect with the data equity principles *Data represents people* and *Data has consequences*, knowing that safety and confidentiality are essential for the populations they serve.

CASE STUDY B

ORGANIZATION B is a larger organization operating in a major urban centre.

90 employees	Major urban centre	Established data practices	Exploring data and equity intersections

They have been using a Client Management Software database for several years and as a result have fairly set data practices, although they do not have an explicit agency-wide policy around data management. Funder reporting and program assessment are their main uses of data, though they believe there is a great deal of untapped potential in the data they collect.

For Organization B, these potential data uses must be balanced against clients' right to privacy. One of their priorities is making sure that everything they collect serves a purpose that ultimately benefits newcomers and their communities. As a result, they connect most strongly with the principle *Data management must balance public good and personal privacy.* Their previous and ongoing work on anti-racism and LGBTQ+ positive spaces has also led to an organizational commitment to the principle *Data is subject to bias,* and one major reason they want to make more use of their data with respect to GBA Plus and equity work is its capacity to challenge assumptions.

Identify data needs & goals

The foundational step to developing a data management strategy is considering data priorities.

This process requires agencies to identify the questions they have which can be answered through data as well as how and with whom they will be sharing data. Non-profit organizations often face a disconnect between their data practices and actual data needs,[5] an issue that leads to gaps in desired data as well as data going unused or underutilized.

USING THIS SECTION

This section will guide you through two methods for identifying your agency's data needs: key questions and data narratives.

Use this section in three ways, depending on what is most useful to you:
1. Read through the text.
2. Reflect on the case studies.
3. Complete the Data Questions and Data Narratives worksheets.

Some parts of the case studies will be more relevant to your work than others—and that's okay! Adapt this process to fit the context and needs of your organization.

Agencies should begin this process by considering what work they want their data to do, including the questions they have about their own work and the data needed to answer those questions. This process benefits greatly from collaboration across teams and departments, as each program's staff may have different questions and hold different pieces of the larger puzzle. Consider creating a small group to take the lead on this process with representation from different teams or programs who can be bridges for questions, needs, and ideas. This group can use Worksheets 1 and 2 to collaboratively build a comprehensive picture of the agency's data needs and goals.

Data equity remains an important aspect of this process, so consider not just what you want to know, but *why* you want to know and who that knowledge will impact, positively or negatively.

To fully understand the ways data is currently used, and to explore how data *could* be used, consider the following questions:

- Who are we serving? How can we tell what groups may be underutilizing our services?
- How do we measure the effectiveness of our work for diverse groups of newcomers?
- What are the narratives we want to be able to tell? Who do we want to reach with them?
- What data is needed for these narratives and measures? What is missing?
- Who is missing from our data? Who is going overlooked or underrepresented?
- Are we using all of our data? Is there data we can better use or stop collecting?

QUICK TIP

As much as possible, this group should also incorporate staff of different identities and backgrounds. When diverse perspectives are both present and valued, they can help locate each other's blind spots and biases. This furthers both GBA Plus *(Challenge assumptions)* and data equity *(Data is subject to bias)*.

As part of this process, reflect on the impacts of data practices. From whom is data collected? How is it used? What is the impact on this use? How can harms be reduced and benefits increased for newcomers? Ponder these questions as you complete Worksheet 1. The following examples may help spur ideas.

CASE STUDY A

ORGANIZATION A decides they want to better understand their client populations, so they identify their first key question as: "Who are we serving?" In considering the role of data in answering this question, they determine that demographic data will be the most useful. They currently collect some of this data (such as age and mother tongue) for funder reporting, but the information they want goes beyond what their funder requires, so they identify this as a data need. But this organization does not want to collect information without a good reason, so they consider how they will use this data and what risks its collection may present to clients.

They also want to be able to evaluate their programs, framing this with the question: "How effective are our services?" They decide that this requires usage statistics, such as program enrollment, as well as the ability to disaggregate demographic data so that they can use GBA Plus to determine how different client populations access and benefit from their services. This will require them to work with their database contractor to build in some reporting templates that will perform the collation and disaggregation they need.

"Who are we serving?"

| How can data answer this question?	What are potential impacts?
• Demographic data such as age, gender, population group, sexual orientation • We collect some of this data	• Ability to offer more targeted programs • Possibility of building connections with other community organizations • Safety concerns for some clients (e.g. LGBTQIA+)
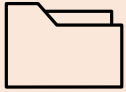
Asking Data Questions

CASE STUDY B

ORGANIZATION B has a slightly different question: "Where are the gaps in our services?" As they reflect on this question, they decide that they also want to know where clients are most successful, so they reframe this question as "How do diverse newcomers experience our services?" Digging into this question, they realize it has several layers. First, they need to know who is accessing their services, which means they need demographic data. Because they have decided to prioritize integrating GBA Plus, they need the ability to disaggregate this data.

Then, they need to know how their services are being used, which they decide will require usage and outcome statistics. To find those statistics, they need several pieces of data, such as enrollment, departure from services, and milestones like finding employment or completing a LINC course. Although they already track raw enrollment numbers, they do not currently link clients' demographic information to their service usage. They identify this linkage as a data need.

Finally, they decide that they want to incorporate clients' personal experiences through a satisfaction survey. This entails a special project, but they value qualitative data for its ability to contextualize their numbers and check their biases.

"How do diverse newcomers experience our services?"

How can data answer this question?	What data do we have?	And what data do we need?
• Demographic data can tell us who is accessing a given service • Enrollment statistics can tell us what services are most and least used • Milestones and outcomes can tell us where clients are having the most and least success • Linking these three can help us find patterns in who is having success and where they are having it • What clients tell us about their personal experiences helps us understanding why some services work better for some clients than others	• Demographic data • Outcomes • Enrollment	• Disaggregation • Milestones • Linkage • Qualitative client experience
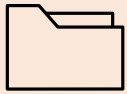
Ultimately, data is a storytelling tool.

The example about LINC student withdrawal on page 5 shows one way that agencies can tell a story using data – in this case, it is a story about parents in their language-learning journey. It is also a story about the organization – how it works with different clients to find the best ways to meet their needs. The same data can be used to tell many different stories, so agencies need to consider who they want to communicate with in order to determine what evidence would be most compelling for that audience. Understanding who is listening to the story and what the story is about shapes the supporting details (i.e. data) that make the story engaging, believable, and convincing. This process can be thought of as creating a *data narrative*. It is a way for the organization to identify what data needs to be collected, how it will need to be organized for the greatest utility, what gaps exist, and to consider not just who each data narrative is for, but also who each data narrative *affects*.

To support GBA Plus integration, organizations need to identify intersectional identity factors to be considered within each data narrative. Upholding data equity means agencies must frame this process through the principles outlined on pages 8-9. In particular, remember that *data has consequences* – the language used in a data narrative can have profound impacts on how its audiences react to the information. One suggestion is to move away from deficit framing; the question asked above by Organization B is an example of this shift, as it looks to client experience to locate successes as well as gaps.

When constructing data narratives, the key pieces are the:

- **PURPOSE** (why you are telling this story and what you hope to achieve)

- **AUDIENCE** (who you will tell this story to)

- **EVIDENCE** (the supporting information that will demonstrate or prove your purpose for the identified audience)

Use Worksheet 2 to explore the data narratives your organization wants to tell. On the next page are two examples of organizations designing data narratives as outlined in Worksheet 2.

QUESTIONS OR STORIES?

You may find as you consider your agency's data needs that data questions are more useful to your process than data narratives, or vice versa. Feel free to use one or both as works best for your agency, team, and context.
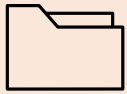
CASE STUDY:
Building Data Narratives

CASE STUDY A

ORGANIZATION A considers the impacts of better understanding the populations they serve, and finds a connection between their original question ("Who are we serving?") and their goal of building stronger connections with other community organizations who serve similar populations (such as ethno-cultural or religious organizations) or populations with similar needs (such as employment or food insecurity). They determine that their purpose is to build community connections; their audience is other community organizations in the area; and the subject of the narrative is who they serve. In order to convince ethnocultural or faith-based organizations that collaboration will benefit their shared clients, they need demographic data that will show the overlap in client populations. To connect with organizations such as food banks, the organization may instead want to look at needs assessments and referrals, as well as case notes about needs that arise throughout the client's journey. They consider the importance of statistical evidence, such as "We hope to 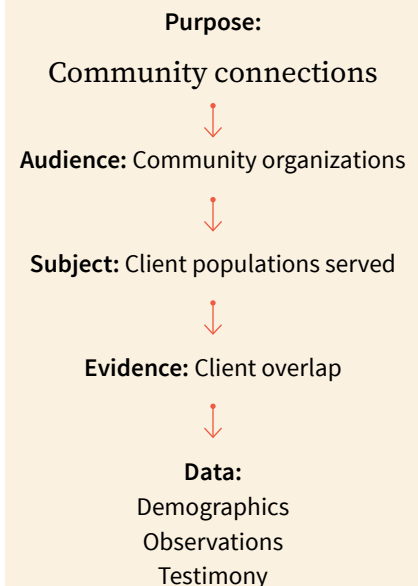build a relationship with the local Filipino cultural association because we currently serve 47 Filipino clients" (which they plan to draw from demographic data) or "We want to connect with local food banks because 32% of our clients have used a food bank in the past six months" (information they intend to discover through case notes). From this process, they identify demographic data and service usage as data gaps, which overlaps with the gaps identified through their question on .

Data Narrative A

Purpose:

Community connections

↓

Audience: Community organizations

↓

Subject: Client populations served

↓

Evidence: Client overlap

↓

Data:
Demographics
Observations
Testimony

CASE STUDY B

ORGANIZATION B knows that they use data to report to funders, but want to deepen their use of this data. They take the subject "Client experiences of our programs" and decide that what they want is to evaluate their programs' effectiveness for different groups of clients. For this, they want to be able to make use of these evaluations internally, to discover where interventions would be most impactful, as well as sharing with funders. To develop findings in this context, they need to be able to locate successes and gaps – both who is benefiting most and why as well as who may be falling through the cracks and why. This analysis will require demographic data (which they will need to disaggregate) as well as service usage, program outcomes, and possibly milestones. They already have or collect much of this data, though they do not currently link the different pieces or have a systematic way to collect client feedback, so they identify both linkage and client feedback as data gaps – this overlaps with the data gaps they identified through their data question on . In the next step, Evaluation, they will take a closer look at what they collect and how it is stored and used.

Data Narrative B

Subject:

Client experiences
of programs

↓

Purpose: Program evaluation

↓

Audience: Internal use | Funders

↓

Evidence: Successes | Gaps

↓

Data:
Demographics | Service usage
Outcomes | Client feedback

1
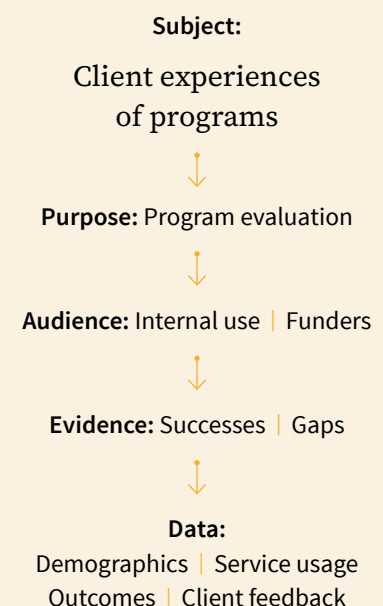
Identifying Data Needs

Use this worksheet to explore ways your organization currently uses data and how you want to use data by developing questions you want to answer through data. This can be completed through group discussion, or it can be completed individually and then shared and discussed as a team.

Key Question:

1A How can we use data to answer this question?	1B What data do we have that can answer this question?	1C What data do we need to answer this question?

Key Question:

1A How can we use data to answer this question?	1B What data do we have that can answer this question?	1C What data do we need to answer this question?

Building Data Narratives

To complete this worksheet, consider who you want to tell about the work you are doing (audience), what you want to share with them (subject), and why you want to share with them (purpose). Then consider what information this audience would value and what data you need to have access to for that purpose. This worksheet can also be completed individually and shared or worked through collaboratively.

2A Purpose:

2B Subject	**2C** Audience(s)	**2D** Evidence needed
•	•	•
•	•	•

2E Data required $\longrightarrow$	**2F** Data available $\longrightarrow$	**2G** Data gaps
•	•	•
•	•	•
•	•	•
•	•	•

2A Purpose:

2B Subject	**2C** Audience(s)	**2D** Evidence needed
•	•	•
•	•	•

2E Data required $\longrightarrow$	**2F** Data available $\longrightarrow$	**2G** Data gaps
•	•	•
•	•	•
•	•	•
•	•	•

Evaluate current practices

Once an agency knows what it wants data to do, it can determine whether its current practices meet those needs.

This section provides an overview of core data practices and an introduction to evaluating them. There are also case studies to provide an example of an evaluation's end result. Visit Appendix A for more detailed descriptions of the core practices and Appendix B for worksheets to guide evaluation of each practice.

Below is a suggested order for working through this content, but you can go through it in the way that best suits your work.

1. Read through the text in this section.
2. Reflect on the case studies.
3. Reflect on the questions in the Equity and Values Evaluation worksheet in this section and add your agency's values.
4. Review Appendix A: Core Practices
5. Complete the worksheets in Appendix B.
6. Return to the Equity and Values worksheet in this section.

The following summary outlines the evaluation process; refer to:

- Appendix A for a more in-depth description of the practices listed below

- Appendix B for worksheets supporting the evaluation of each of those practices

- Worksheet 3 in this section to frame your evaluation – place it prominently in the space where you conduct the evaluation

- As you complete Worksheets A1-A6, return to Worksheet 3

Consider how your current practices embody these values:

- What agency values do you most want to incorporate into your data work?

- What data equity principles resonate most strongly with your organization?

- What changes might bring your practices more in line with these principles?

Remember these questions as you examine the following core data practices.

- **STANDARDS:** The rules used to standardize how data is described, represented, and structured. Agencies should consider how their data standards align with those of collaborators, such as the Government of Canada, to enable interoperability and comparative analysis. Consider whether categories and labels align with the experiences and self-identification of communities from whom data is collected.

- **ANALYSIS AND USE:** Access to disaggregation and collation, how information is presented and distributed, and the context in which analysis takes place. Key to this process is examining how the use of data impacts newcomers and whether it empowers or silences their communities. Considerations include ensuring fair and accurate reporting that engages with the context in which data is situated; monitoring for bias; and the care with which analysis is framed and presented. This process benefits from direct engagement with client communities, as qualitative data can be used to contextualize quantitative analysis, including checking to see if data analysis aligns with lived experience.

- **QUALITY:** The accuracy, completeness, and consistency of data, which ensures data is trustworthy and provides a strong evidence base for decision-making and data narratives. Agencies should establish a clear internal target for data quality and practices for managing inaccurate, incomplete, or inconsistent data.

- **COLLECTION:** What is collected, how and when the collection takes place, and the establishment of culturally safe spaces for data collection. Informed consent requires that staff be able to explain why they are collecting data and what it will be used for. It is also important to understand power relations and the sensitivity of personal data in order to establish culturally safe collection practices. There are many methods for collecting data, ranging from direct questions (as in a client intake) to observation (for example, by settlement workers who then record their observations in case notes), to focus groups and surveys. Note that surveys can produce quantitative data through scaled questions (e.g. "Rate your experience with this program on a scale from 1 to 5") and qualitative data through open-ended questions. Each method gives slightly different data, provides different context to that data, and varies in accessibility (for example, surveys rely on literacy in the language used), so consider who you want to reach, exactly what you want to know, and what methods are most inclusive for that group of people. Remember that accessibility results in more accurate data.

COLLECTION METHODS

There are many ways to collect data. As you decide what data your organization needs, consider how best to collect it. In 1:1 interviews, such as client intakes? Through observations recorded via case notes? Via scaled response or open-ended surveys? Focus groups? Whatever method you choose, be intentional about accessibility and think about who might get left out.

- **CURATION:** Where and how data is stored, the protections in place for client privacy, database structures, cleaning and processing data for use, and metadata. Both privacy legislation and contract stipulations may place requirements for how data is stored and protected. Another aspect of curation is how databases are structured; it is recommended that agencies establish database structures that enable disaggregation based on the factors identified in the previous stage.

- **MONITORING AND EVALUATION:** Ongoing monitoring and regular evaluation ensures data practices are consistent and continue to meet standards and needs over time. This may include regular audits for data quality, reviewing curation and privacy protections, and providing training and resources to staff responsible for collecting data, among many other possibilities. GBA Plus needs, agency values, and data equity principles need to be built into any monitoring and evaluation framework.

DATA BIOGRAPHIES

Metadata is basically the story of how data came to be. Much like a biography of a person, or a news article covering an event, metadata describes the Who/What/Where/When/Why/How of the data's collection and processing. This provides subsequent users with context that helps them understand the limitations, possibilities, and ethical implications of a collection of numbers.

It is productive to have multiple people involved in this process. This is both because the information required to conduct a thorough evaluation is rarely held by only one person, but also because multiple viewpoints help to counteract the biases individuals cannot see on their own.

CONSIDERATIONS FOR DEMOGRAPHIC DATA

In the previous section, both organizations in the case studies identified demographic data as key to their purposes. It is important to note that demographic data requires special consideration in its collection, curation, and use. More than any other kind of data, demographic data directly represents the people from whom it is gathered. It is information that can be linked directly to specific people unless it is anonymized and aggregated with the information of other individuals. This is one reason some analysis omits certain categories – if there are too few people of a particular identity, it is much easier to identify individuals, stripping from them the protection of anonymity.
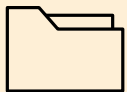
Evaluating Data Practices

CASE STUDY A

ORGANIZATION A brings together their staff to discuss their current data practices. During this evaluation, they find several opportunities to build the practices they want through a data management strategy.

They currently collect limited demographic data, primarily for funder reporting, so they have no internally set standards beyond those required for reporting. This means they will be able to consult with newcomers and staff to set updated internal standards that represent the clients they serve. They also receive regular reports from IRCC with statistics about their work. To ensure that they will be able to compare their internal data to external data (such as reports from IRCC or Statistics Canada), they want to balance those consultations with standards used by the Government of Canada, information that is available on the Statistics Canada website. This also means that they can work with their database contractor to design data fields (rather than having to retrofit existing fields), and they will be able to set up quality targets that best represent their goals.
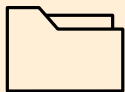
When looking at their data collection practices, they find a pivotal strength – staff are experienced and skilled at building rapport and trust, which they can draw on to create guidelines for collecting data. For example, staff's strong relationships with their clients reveals that clients feel especially hesitant to share information about their sexual orientation and/or gender-expansive identities. Even when clients feel safe disclosing that information to staff, sometimes they request that staff not record that information. This means that work will need to be done to create an environment where clients feel confident that their information is safe with the organization.

This understanding leads them to highlight *data management must balance public good and personal privacy* when they return to Worksheet 3 – while it would be beneficial for the organization to know this information about their clients, they need to balance that with respect for clients' rights to make decisions about their own data. This ties in with one of their agency's values – respecting the autonomy and personal journeys of newcomers – which they also record on Worksheet 3.

CULTURAL SAFETY

The Government of Canada Public Health Authority defines cultural safety as "an outcome based on respectful engagement that recognizes and strives to address power imbalances" that are an innate part of a system.

This concept arose from interactions between Indigenous New Zealanders and the healthcare system, but it is relevant in any contact between a minoritized person and a larger system. There are a number of Indigenous-led organizations in Canada who provide cultural safety trainings for a variety of contexts.

Evaluating Data Practices

Standards	Collection	Quality
• Few set internally • Do receive statistics from IRCC using Government of Canada standards	• Excellent rapport and relationships with clients • Strong organizational commitment to cultural competence and safety but no formal trainings • Currently collect limited demographic data, but database can support more robust data • Do have case notes field in database which can be used to track client needs and progress	• Do not currently conduct data audits • Are well-situated to set quality targets

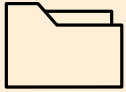| Curation	Analysis	Monitoring
• Database meets privacy requirements • Strong relationship with contractor • Desire ability to perform disaggregation	• Currently limited by available data • Strong established relationships with newcomers and communities	• No structures currently in place • Ability to build in practices as part of strategy
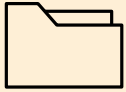
Evaluating Data Practices

CASE STUDY B

ORGANIZATION B finds, when they evaluate their data practices, that they already collect a lot of the data they want. The demographic data they currently collect does not quite match what they need to meet their goals, so they will need to bring those into alignment.

More importantly, they find that each team or department has slightly different practices with respect to collecting, entering, and using data. This is problematic for comparing data across the organization, but it does give them the opportunity to create guidelines that build on the best practices arising from a variety of experiences with data within the context of their organization. They also benefit from a staff composition that is a reasonable reflection of their current understanding of the diversity of the newcomer communities in their area. They note this on Worksheet 3 as a data equity measure supporting the principle: *Data is subject to bias.*

Consultation with their database contractor uncovers capacity within the database structure to develop functions that meet their data needs. For example, although they do not currently audit their data for accuracy and completeness, their database has untapped capacity to do so. Exploring this facet of data work resonates with their agency value of excellence, which they enter on Worksheet 3.

When looking at the ways they currently analyze their data, they find that regular program evaluation is an established practice, but that a lack of disaggregated data means that they have not been able to fully support GBA Plus intersectional analysis, an issue which they record on Worksheet 3. On the other hand, they also realize that their active and ongoing engagement with newcomers and their communities presents opportunities to consult on self-identification labels (for data collection and curation) as well as learning about questions clients have that agency data work could answer, and they make a note to consider engagement methods in the next phase: designing their data management strategy.

Evaluating Data Practices

Standards	Collection	Quality
• Inconsistent across teams • Inconsistency limits interoperability • Demographic data does not fully support data goals	• Already collect much of the desired data • Staff composition reasonably reflects populations served • No standard training on collection practices • Do not currently track milestones	• Do not currently check for accuracy or completeness • Discover this capacity is available within database

Curation	Analysis	Monitoring
• Database meets privacy requirements • Allows disaggregation by demographics, but not collation with program statistics • Strong anonymization capacity	• Regular program evaluation well-established but lacks disaggregation • Active, ongoing engagement with newcomer communities presents opportunity for consultation	• Each team manages their own data checks • No agency-wide structures currently in place

3

Reflecting on GBA Plus and Equity

This worksheet provides questions for reflection. Consider posting these questions prominently as you complete the worksheets in Appendix B. As part of this process, consider how you want to engage in GBA Plus, where your agency's values are supported or contradicted by existing data practices, and which data equity principles are most relevant to your work.

3A   Assessing GBA Plus Data Capacity

Identify Issues
What structures do we have to support GBA Plus analysis?

What disaggregation are we able to perform?

How are we tracking and defining success for our clients?

Challenge Assumptions
What identity factors are we seeing among our clients?

What do we think we know about their impact on client experiences?

What does the data show?

Gather the Facts
What analysis are we doing?

What analysis do we need to answer our key questions?

What data do we have and where are the gaps?

3B   How do our data practices uphold our values as an agency?

Agency Value:

Agency Value:

Agency Value:

3C Assessing Data Equity Measurements

Data represents people.
How are we ensuring that our strategy treats our clients and their data with respect?

Data is subject to bias.
How are we guarding against bias at each stage of the data lifecycle?

Data has consequences.
How are we ensuring fair and accurate representation of our clients?

Data does not exist in a vacuum.
How are we considering context in our data work?

Data must balance public good with personal privacy.
How are we balancing the needs of the agency and the needs of newcomers?

Data management is a learning process.
What accountability measures do we have in place?

Design data management strategy

Once agencies understand what data they need, the purpose it fulfills, and how current data practices align with these needs and priorities, they are ready to design a data management strategy.

- [Worksheet 4](#) provides a template for designing data practices that meet the data needs identified through Worksheets 1 and 2.

- [Worksheet 5](#) draws from Worksheet 3 to articulate the role GBA Plus, agency values, and data equity principles play in your strategy.

During the design phase, use these questions to frame the transition from identifying needs and evaluating practices to developing policies and practices.

- Where do existing data practices align with GBA Plus and data equity principles? Where can this alignment be strengthened?

- What strengths can the agency use to improve data equity and support intersectional analysis?

- What steps are necessary to bring data practices in line with data needs?

- What strengths can the agency deploy to address data gaps?

- What resources are available for this process? What resources (funding, training, etc.) will be needed to make these changes?

Hiring practices that support diverse staff can help to ensure a variety of perspectives on data within the organization, as can direct engagement. Diversity among newcomers means they cannot be treated as a monolithic group. Agencies should consider ways to engage with groups, such as newcomers with disabilities; from different ethnic, cultural, linguistic, and religious backgrounds; of different immigration statuses and streams; LGBTQIA+ newcomers; and other groups, remembering that differences exist within groups as well as between them. When designing engagement strategies, consider how data represents, silences, and empowers communities.

USING THIS SECTION

In this section, you will bring together the results of the previous sections—Identifying Data Needs and Evaluating Data Practices—to design a data management strategy for your organization.

Use this section in three ways:
1. Read through the text.
2. Reflect on the case studies.
3. Complete the accompanying Worksheets using the data gaps from [Worksheet 1](#) and [Worksheet 2](#) and the agency values and data equity principles identified through [Worksheet 3](#).

Remember to consider your GBA Plus data goals from [Worksheet 3](#) as you design your data management strategy.

Your strategy should include:

- A description of the principles and values that govern your agency's data work;

- An overview of the purposes data fulfills within the agency, including a section on how GBA Plus analysis is supported and managed;

- Internal guidelines for equitable data management through core practices;

- Training and resources to support those guidelines; and

- Accountability measures, including ongoing monitoring and evaluation procedures.

- Logistics, including budgetary requirements, may also be included.
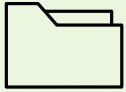
CASE STUDY A

ORGANIZATION A compares their identified data needs (Worksheet 1 and Worksheet 2) to the evaluation they conducted in the previous phase (Worksheets 3 and A1-A6). The most prominent need they have identified is strengthening their disaggregable demographic data. To meet that need, they know that they need two things: demographic data (which will need to be collected directly from clients) and database structures enabling disaggregation. They identify staff who are invested in this process and committed to equity in their work; the strong relationships that exist between clients and the organization; and their connections with the database contractor as key resources. They decide that they need to update their existing data standards (including labels and categories) and build in new standards for the data they will begin to collect, design database report templates with the support of their database contractor, and develop staff training on data collection and cultural safety. To support these action steps, and to keep them grounded, they decide to begin by connecting with clients and staff to establish data standards that align with how clients self-identify – for example, with respect to ethnocultural identities.

THE RIGHT DATABASE

There is no One Right Database. Selecting the right database or CRM software for your agency is a complex process balancing cost and ease of use against the capacity to provide the functions required to meet your data needs. That is beyond the scope of this guide—just remember that fully engaging with GBA Plus requires the ability to track demographic data and usage statistics, and to disaggregate that data for analysis.

Data need ⟶	Data gap ⟶	Resources available ⟶	Action steps
Recorded on Worksheet 1 • Disaggregable demographic data	• Demographic data • Database structures for disaggregation	• Engaged staff • Strong client relationships • Database contractor relationship	• Client engagement • Set data standards • Design database report templates • Develop staff training
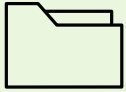
Designing a Strategy

CASE STUDY B

ORGANIZATION B evaluates their existing data practices, and finds that the quality of their data was not where they wanted it to be. They already collect most of the data they want, but variations in standards and practices limit their ability to check data quality. To do this, they pull together the various teams and departments to build a best practices data collection and standards guide drawing on the successes and challenges each team has encountered. They also set a data quality goal of 95% complete data in their database, and work on a process for conducting a regular audit of data in their possession.
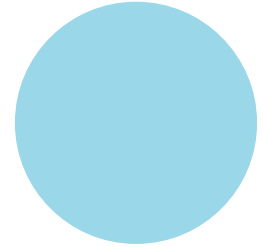
Data practices $\longrightarrow$	Goals $\longrightarrow$	Resources available $\longrightarrow$	Action steps
From Worksheets A3 & A6. • Data quality • Monitoring	• Consistent standards • Quality metrics • Regular data audits	• Staff knowledge • Statistics Canada data literacy resources	• Develop best practices guide • Set data quality goal • Establish audit practices and schedule
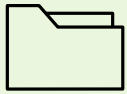
CASE STUDY A

ORGANIZATION A looks back at the principles of data equity they explored through Worksheet 3 as part of developing their strategy. The concerns of LGBTQIA+ newcomers, and their fears of being outed in a small community, have stuck with the team. This understanding leads them to connect most strongly with two of the principles from pages 8-9: *Data represents people* and *Data has consequences*. They decide that there are two sides to respecting client identities and the data associated with them – accurately recording the information clients share, and clearly understanding the limits and boundaries clients set around the use of their personal information. For the first part, their related activities include adding fields to client profiles so that preferred names and pronouns are prominently displayed. For the second, they want to work with their contractor to add an equally prominent field where staff can record when to use the name and pronouns (e.g. only in one-to-one meetings, all the time, etc.) as well as engage staff in training around informed consent and cultural safety.

Example A

Equity principles
- Data represents people
- Data has consequences

↓

Alignment
- Importance of self-identification
- Client consent and boundaries

↓

Activities
- Add client name and pronouns to profiles
- Add field to note disclosure and use limitations
- Staff training on informed consent and cultural safety

CASE STUDY B

ORGANIZATION B also refers to Worksheet 3 as they look at their data narratives from a GBA Plus perspective (see pages 6-7), clearly outlining the relationship between GBA Plus and their analysis goals: they want to use disaggregated data to Challenge their assumptions about program use and help identify gaps, with the goal of creating better outcomes for diverse newcomers in their city. To support this, they feel strongly that they need to integrate another data equity principle from page 9 – *Data does not exist in a vacuum*. Because their data goals include program evaluation, they want to make sure that they are taking into account context that may not appear in the numbers. To do so, they develop a client satisfaction survey; deciding that it would be beneficial to conduct this survey on a regular basis, they document their process in order to build a set of guidelines for future surveys.

Example B

Equity principles
- Data is subject to bias
- Data does not exist in a vacuum
- Challenge assumptions

↓

Alignment
- Active client engagement
- Misalignment of practices with GBA Plus

↓

Activities
- Establish disaggregation processes
- Incorporate client satisfaction survey

4 | Designing for Data Needs

Refer to Worksheet 1 and Worksheet 2 to fill in the Data Need boxes below. Determine the method for collection (such as client intake interviews, surveys, etc.). Consider the resources your organization has available to you—funding, time, expertise, relationships, partnerships, etc.—and formulate action steps to address the gaps. You may end up using multiple copies of this worksheet. This process benefits from collaboration.

4A Data Need	**4A** Data Need	**4A** Data Need
(From Worksheet 1C)	*(From Worksheet 1C)*	*(From Worksheet 2E)*
4B Collection Method(s)	**4B** Collection Method(s)	**4B** Collection Method(s)
↓	↓	↓
4C Resources available to address need	**4C** Resources available to address need	**4C** Resources available to address need
↓	↓	↓
4D Action Steps	**4D** Action Steps	**4D** Action Steps

Designing for Equity

To complete this worksheet, use the results of your reflections as recorded on Worksheet 3. Carry over the agency values identified in section 3B and identify the data equity principles from section 3C that resonate most with your organization's work and culture. Consider how your data practices can best reflect those values while still meeting the data needs described on Worksheet 4. Be strategic about the specific values and principles you want to address. You may need more than one copy of this worksheet.

5A **GBA Plus**
(From Worksheet 3 Section A)

What structures are we implementing to support GBA Plus analysis?

What types of analysis are we conducting?

What identity factors are we accounting for?

⟶ **5D** **How is this integrated into our strategy?**

5B **Agency Value:**
(From Worksheet 3 Section B)
Example: Integrity

⟶ **5E** **How is this integrated into our strategy?**

(Example: We hold ourselves to a high standard when protecting client privacy and confidentiality. We have set accountability measures to mitigate possible bias and produce analysis that is empowering rather than harmful.)

5C **Data Equity Principle:**
(From Worksheet 3 Section C)
Example: Data is subject to bias.

⟶ **5F** **How is this integrated into our strategy?**

(Example: We base our standards on consultation with clients, train our staff to collect data in a non-judgmental manner, and filter our analysis through a collaborative team of diverse backgrounds.)

Implement data strategy

The revision of existing policies and introduction of new ones is a process. Implementing a data management strategy may entail significant changes to how the agency collects, stores, and uses data.

Ground this process in the desired organizational changes and in the identified data needs and goals. Remember that developing consistency requires more than writing out a manual: it demands the integration of data equity principles, including engaging with staff to ensure consistent understanding of the data management strategy and its purpose. Implementation also means learning how the data management strategy, once put into practice, achieves or misses the desired results, including how it affects, benefits, or hinders the work of staff and the experiences of clients.

Because it can be difficult to anticipate how policies will impact diverse groups, it is essential that the implementation stage be treated as flexible. Worksheet 6 provides some guiding questions, including incorporating a process for checking in with staff and clients to identify any unintended consequences, mechanisms for both staff and clients to provide anonymous feedback, and a plan for how to follow up on concerns for further details and possibilities for addressing these issues. A designated strategy lead, team, or committee may be best positioned to follow through with investigating and proposing adjustments to the strategy.

The necessary program and policy developments and interventions may need to be implemented over time. Use Worksheet 7 to develop a timeline with appropriate steps, including time to test and evaluate new tools, procedures, and policies. In accordance with the design phase, ensure that the proper tools and training are provided to staff. Investigate what tools and training are available that align with the agency's strategy and goals. Consider cross-agency collaboration if other organizations in the sector are also developing data management strategies in order to share the resource burden of developing tools and training. While it is important to consider the role of each practice, organizational priorities and/or resource limitations may require agencies to focus on a few key practices at a time.

USING THIS SECTION

Throughout this section, you will be asked to consider both the logistics of implementing the strategy you designed in the previous section as well as equity concerns for the manner in which you enact these practices.

Use this section in three ways:
1. Read the explanatory text.
2. Reflect on the case studies.
3. Complete the Implementing Data Practices worksheets for the core practices you are focusing on.

Your organization might focus on some, most, or all of the practices described in Phase 2 depending on your needs and resources.

Once the data management strategy has been tested, proved to uphold data equity and achieved the desired results, agencies can move from this stage to ongoing monitoring and evaluation.
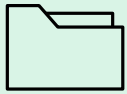
CASE STUDY A

ORGANIZATION A begins to put their strategy into practice by connecting with clients to inform the standards for the new data they will be collecting. While they want to be able to compare their data with Government of Canada datasets (such as IRCC and Statistics Canada), they also want to make sure that the categories and labels they use accurately reflect how their clients self-identify. They do this both through informal conversations between settlement workers and clients as well as through social events – at a semi-regular cultural event, they set aside some time to talk about data with the newcomers who are attending.

Once they have the categories and labels they intend to use, they begin training staff on the newly designed protocols for data collection and entry. They test out the changes made to their database. As they work through this process, they realize that it would be beneficial to create a living data guide for current and future staff that goes through the new practices, why they exist, and also orients new database users. As they continue to work with the database, they add new knowledge to this document. They also set up a schedule for when they run reports as well as two ways for staff and clients to leave anonymous feedback: a box for written comments in a shared space, and an online survey through Google Forms which people can complete as many times as they want.

Policies	Practices	Trainings	Accountability
• Demographic data collection at intake • Entry into database along with needs assessment	• Demographic data considerations • Informed consent • Database processes	• Cultural safety • Privacy and consent • Data standards • Database use	• Comment box and Google Form • Staff meeting discussions • Client engagement fora
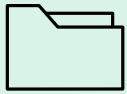
CASE STUDY B

ORGANIZATION B develops a standardized intake form based on the best practices from each team and department, which they distribute across the agency. As part of this rollout, they use their quarterly all-staff meeting to conduct a training session to make sure everyone is familiar with the changes. Prior to this training, they encourage staff to watch videos from Statistics Canada's data literacy video series.

Once these new practices are in place, they give staff two months to adjust to the changes before launching their first data audit as part of the strategy. Over the next six months, they begin to see an increase in the completeness and consistency of their data. Using these audits, they also notice patterns in which fields are most frequently unanswered, and which fields often have write-in labels. They can use this information to explore the context for these responses, which will allow them to update their standards and/or provide targeted support to staff. To keep their journey-mapping project moving forward, they expand their database report templates to support disaggregation and begin pulling these reports to analyze.

Policies	Practices	Trainings	Accountability
• Consistent data standards across teams • Regular audits • Data quality procedures	• Standardized intake form • Data quality targets • Management of missing, incomplete, or invalid data	• New intake training • Statistics Canada data literacy series	• Data audits • Training completion • Training feedback survey

6

Implementation Planning

Use this worksheet to systematically outline the processes for implementing the data strategy designed through Worksheet 4 and Worksheet 5. Consider logistical questions as well as how this implementation can be done in alignment with principles of equity and agency values.

6A Logistics

Who is taking the lead on this process?

Who is being consulted in this process?

What core practices are we focusing on?

How are we defining success for our data management strategy?

How are we measuring this success?

How are we checking in with our staff about the effects of this strategy on their work?

6B Equity

What data equity principles are we prioritizing?

How are we integrating those principles?

How are we assessing equity?

What identity factors have been taken into account (e.g. race, immigration status, etc)?

Are the policies, guidelines, and/or trainings provided to staff accessible and inclusive?

How can staff and clients provide anonymous feedback during this process?

What mechanisms do we have for following up on concerns?

This worksheet is best completed after completing Appendix B. Take the action steps from those worksheets and from Worksheets 4D, 5D-F, and 6A-B to lay out how your agency will work with data, using the core practices described in Phase 2 and in Appendix A.

7A Core Practice:

7B Policies

-
-
-

7C Practices

-
-
-

7D Training & Resources

-
-
-

7E Accountability Measures

-
-
-

7A Core Practice:

7B Policies

-
-
-

7C Practices

-
-
-

7D Training & Resources

-
-
-

7E Accountability Measures

-
-
-

Monitor and evaluate data strategy

The final stage of data management is the establishment of monitoring and evaluation mechanisms.

At this point, the organization moves from the iterative process of implementation and refinement to ongoing monitoring and periodic evaluation as laid out in the data management strategy.

There is no one set method for managing this aspect of data management, but it is important to frame the associated processes using data equity principles and GBA Plus considerations. This process can incorporate and build on the data evaluation stage explored earlier in this guide, but it can also be developed in line with existing agency measures for monitoring and evaluation.

USING THIS SECTION

At this point, you have a data management strategy designed and in place. In this section, you will consider its impacts to gauge its effectiveness and to set up a system enabling future evaluations.

In this section you can:
1. Read the explanatory text.
2. Reflect on the case studies.
3. Complete the Monitoring Data Equity worksheets.

These worksheets ask you to look critically at your data practices (feel free to reference the Evaluating Data Practices worksheets in Appendix B) both logistically and in terms of equity.

Some key questions that should be considered include:

- Who is checking for alignment with data equity and GBA Plus? How are we integrating and amplifying diverse perspectives into this process?

- How are our practices upholding principles of data equity?

- What intersectional analysis are we supporting with our data? Are there gaps?

- Are we able to support the data narratives we want to tell with adequate evidence?

- Are our practices in line with the most up-to-date legislation, contract stipulations, and commitments we have made?

- Does our data meet the quality standards we set for ourselves?

- Where are we excelling at data work, and where are our areas for growth?

- Are our data practices responsive to our needs as they develop and change?

- Are we checking qualitative data from newcomers to ensure our quantitative data is consistent with their lived experiences? How are we investigating incongruities?

- How is our data work empowering or disempowering the newcomers we serve?

This process may include internal audits for data quality, regular training and refresher opportunities for staff, and consultation with newcomer communities to ensure the agency's data reflects the experiences and needs of diverse groups. It must include consideration of impacts and gaps across diverse groups, as it is possible for data work to empower some groups while silencing others.

Finally, ensure there is a process in place for the results of monitoring and evaluation to result in timely adjustments to data policies and practices. An organization's data priorities may change over time, and it is important for practices to reflect those shifts. As well, equity is a constantly negotiated process and the activities needed to support it will change over time. An equity-driven data management strategy needs built-in flexibility that supports these adaptations.
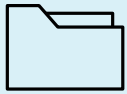
Community Impact & Agency Practices

CASE STUDY A

ORGANIZATION A is scheduling regular open forums for clients to build relationships and have open conversations about equity. In addition, they establish data as a regular agenda item at their recurring staff meetings. Part of these forums is presenting the data analysis that was conducted over the past quarter, including changes in program usage, enrollment, withdrawal, or in client demographics – for example, connecting with staff about increases in clients from certain areas or of certain identities. At one point, they notice they are seeing a noticeable increase in clients from Uganda, so they decide to bring these clients together and talk to them about their experiences and needs. Staff are encouraged to pay close attention during data collection and to make note of any questions that seem to make clients uncomfortable with an eye toward adjusting data collection through label changes or training and environmental shifts (such as toward or away from enclosed offices).

Example A: Community Impact

Whose voices are we hearing?
Who might be going overlooked?
Bringing together clients and staff for conversations about equity; data is a regular topic. Not everyone is able to attend these events, so staff also make an effort to connect and share with clients.

Where and how is our data work empowering the communities we serve?
Connecting with community organizations to share data gives them more tools to further their own work. Providing data directly to clients puts some of the control back into their hands.

Where and how is our data work silencing or disempowering the communities we serve?
There is some concern that focusing on the equity forum events misses clients—both ones with more privilege and ones with the least.

CASE STUDY B

ORGANIZATION B works with each department and team to create methods for staff to engage with clients about data work. They decide that they want data equity to be an ongoing commitment, and the organization's Equity Committee adds data to their area of concern. They also set a schedule for data audits (to assess data quality) and for a regular Client Feedback Survey, which allows them to check the quantitative data in their database against the qualitative data produced by client testimonies about their experiences. These measures allow them to see how their data supports the equity-driven work in which they are engaged as well as checking how equitable their data processes are. These practices are connected to their data equity commitments from Worksheet 5 – *Data is subject to bias* and *Data does not exist in a vacuum*. This also helps them keep GBA Plus at the center of their work, as they challenge their assumptions and consistently consult with the newcomers they serve.

Example B: Agency Practices

Who is responsible for checking alignment with data equity & GBA Plus?
Equity Committee has taken on data as part of its portfolio, both how data works for equity and how equity plays out in data practices.

What role do diverse perspectives play in this process?
Each team or department is represented on the Equity Committee; the selection process includes diversity guidelines.

Where are we checking our quantitative data against the lived experiences of our clients?
The Client Feedback Survey is instituted as an annual practice.

8

Monitoring Data Equity

Once a data management strategy is in place, it must be monitored for both effectiveness and for adherence to the agency's commitment to equity. This sheet focuses closely on equity; use the questions as a starting point when you consider how your work can be assessed for adherence to equity principles on an ongoing basis. Draw on Worksheet 5 and Worksheet 6 to develop mechanisms for investigating answers to these questions.

8A Community Impact

Who is responsible for checking alignment with data equity & GBA Plus?

What role do diverse perspectives play in this process?

Where are we checking our quantitative data against the lived experiences of our clients?

How are we supporting staff in their work with data? (e.g. ensuring our policies and practices are accessible, responding to feedback and concerns, etc.)

8B Agency Practices

To which groups and stakeholders are we accountable?

Whose voices are we hearing? Who might be going overlooked?

Where and how is our data work empowering the communities we serve?

Where and how is our data work silencing or disempowering the communities we serve?

8C Where are we excelling?

8D Where can we grow?

8E How are we responding?

While the previous worksheet focuses most closely on equity, this worksheet is where you will consider logistical concerns. The worksheets from Appendix B may provide a helpful framework for answering the questions below on an ongoing basis. As part of this process, set a regular schedule for evaluation.

9A How well do our data practices...

Meet our internal quality standards?

Provide what we need to make strategic, data-driven decisions?

Support our data narratives with the right data for the intended audience?

Enable the intersectional analysis we need?

Respond to our changing needs and the needs of our clients?

9B How do our practices align with...

Privacy legislation?

Contract stipulations?

Agency values?

Accountability standards?

Data Equity principles?

GBA Plus requirements?

Internal frameworks, such as IDEA?

9C Where are we excelling?

9D Where can we grow?

9E How are we responding?

CONCLUSION:
Weaving it all together

This final section provides suggestions and a sample template for drafting a comprehensive data management strategy.

The template provided here is only one way to organize a data strategy. Agencies may also choose to create a strategy following the outline established through completing the worksheets in previous sections, or to use an internal process and organization method that better fits their context.

When bringing these pieces together, consider including the following sections:

PRINCIPLES: The values under which the agency works with the data of its clients, drawn from the answers generated for Worksheet 3 and Worksheet 5. This could include a description of the agency values identified on these worksheets; any data equity principles that resonate with the organization's work, which may draw on the language and concepts on pages 8-9 of this guide; and the organization's particular commitment to GBA Plus to tailor services for the specific needs and barriers of diverse newcomers.

PURPOSES: Use the questions and narratives created through Worksheet 1 and Worksheet 2 to describe the uses to which data will be put identified through those worksheets.

- This section can also include information from Worksheet 7 if Analysis and Use was one of the core practices for which your agency completed this worksheet.

DATA TO BE COLLECTED: Draw from Worksheet 4 to list the pieces of data needed to meet the data needs described under Purposes. Using this section of the data management strategy to precisely track what data needs to be collected provides a consolidated place where data practices can be reviewed and updated in accordance with changes in agency priorities and data needs.

- Consider listing specific categories of demographic data to be collected, such as age, gender identity, population group, and religion, as well as whether information about program enrollment, outcomes, milestones, and/or client satisfaction will be used. Alongside this information, include the methods by which this data will be gathered, such as during client intakes, from case notes, through surveys, and any other collection methods to be used.

- Consider determining and including the intervals at which this data will be collected – for instance, only at intake (and whether that data can or will be updated if information changes), through annual surveys, etc.

CONCLUSION:
Weaving it all together

PRIVACY PROTECTIONS: Use the notes from Worksheet A4 and Worksheet 9 to outline the measures in place to protect client privacy, including the intervals at which measures will be reviewed and updated. Work out specifics with the vendor responsible for the client database; it can be beneficial to include information such as the vendor, contact information, and process for reviewing and updating privacy protections in this section.

- Information about how the organization handles informed consent, including the mechanism by which clients can withdraw their consent at a later date, can be outlined here or in the next section on policies and practices.

POLICIES & PRACTICES: This section outlines specific policies developed through the data management development process. It is used to describe how you will engage in the various stages of the data lifecycle.

- Use Worksheet 7 for the policies and practices associated with core practices, such as how data metadata is handled. The worksheets from Appendix B may also be useful. Practices set aside to evaluate at a later date can be mentioned in brief here.

- Consider including in this section specific language connecting these practices to the principles outlined in the section on principles at the beginning of the strategy – for instance, the method(s) used for intersectional analysis as outlined under GBA Plus on Worksheet 5, or how bias is checked and mitigated through practices like data collection or analysis.

TRAINING & RESOURCES: In this section, describe the training and/or resources available to staff. This may include requiring staff to complete online trainings, such as the Statistics Canada Data Literacy Series; internal trainings on the policies and practices outlined in the previous section, such as data collection guidelines; and documents such as this strategy or a user guide to the database or CRM software in use at the agency.

- Use Worksheet 7 for the training and resources developed through this process.

MONITORING & EVALUATION: Here you will lay out the accountability measures you have developed. This may include information about how data quality is assessed, how practices are monitored for equity, and how the agency will engage with clients, stakeholders, and communities to ensure accountability. This should include targets for data quality (most often expressed as a percentage, e.g. "All client file data is at least 85% complete with 90% accuracy") as well as a schedule for when audits occur and any remediation procedures if the audits uncover a gap between data quality and targets.

- Worksheet 8 and Worksheet 9 should provide the processes you developed for which data practices are monitored and evaluated on an ongoing basis and how that is done.

AMENDMENT AND IMPLEMENTATION PROCEDURES: For this section, you will want to describe when and how this strategy is amended. This may include things like a procedure for regular updates of data standards through consultation with clients, avenues for staff to express challenges and concerns and a method for updating policies, training, or resources that will improve staff's ability to work with data, and any other guidelines that facilitate this process.

- Worksheet 6 can be used for specific measures, such as ways staff and clients can provide feedback, including anonymous feedback, and how the agency follows up on and addresses these concerns.

- Worksheet 8 and Worksheet 9, on monitoring equity and practices, also provide valuable information, as you may include some of the questions and answers from these worksheets to describe your amendment procedures.

WORKSHEET 10:
Sample Template for a Data Management Strategy

10

This sample template is intended to provide inspiration and guidance as you develop your own strategy. Each section below includes a reference to a specific worksheet section so that you can make use of the work you've done throughout your use of this guide. This template is not intended to supersede existing processes within your agency, only to provide a starting point.

Begin by stating the principles and values that govern how you work with data.
This explains why the practices in this strategy are important and provides a basis for accountability.

GBA Plus	**Agency values**	**Data Equity**
(Draw from Worksheet 5A)	(Draw from Worksheet 5B)	(Selected principles from Worksheet 5C)
·	·	·
·	·	·
·	·	·

Clearly lay out the purposes for your data work; why do you collect data? How do you use it?
(Draw questions and narratives from Worksheet 1A and 2A and/or data needs from Worksheet 4A)

E.g. Funder reporting (May lump funders into a single category or list individually)	*E.g. To support GBA Plus intersectional analysis enabling tailored services responsive to the unique needs of diverse newcomers*	**Other purposes and narratives**
·		·
·	·	·

List the data you collect based on the purposes identified above. Consider referring to data equity principles for possible language about why collection is limited to these areas. (Draw from Worksheet 4A)		**Describe the measures taken to ensure client privacy.** (Draw from Worksheet 9B and develop in coordination with IT and/or database vendor)
E.g. We collect the following demographic information (list factors such as age and gender)	**Other data collected and its method, such as satisfaction surveys.**	·
		·
·	·	·
·	·	·

Outline the specific policies and guidelines that explain how you meet your data needs while upholding the values and principles listed in the first section:

Equity-focused policies (Worksheet 5D-5F):

-
-
-
-

Core Practices (Worksheet 7A-7C):

-
-
-
-

Describe the training and resources available to staff, internally or externally, to ensure they have the support they need to follow the policies and guidelines in the previous section:
(Draw from Worksheet 7D)

Internal Training:

-
-
-
-

External Training:

-
-
-
-

Resources:

-
-
-
-

Lay out the accountability measures used to ensure your organization upholds its data values.
(Use Worksheets 8A-B and 9A-B; may include equity integration from Worksheet 5D-F)

-
-
-
-

Describe mechanisms in place for revising this strategy at need, including ways to receive feedback from staff and clients.
(Draw relevant language and concepts from Worksheet 9)

-
-
-
-

Below are sample sections of the data management strategies developed by our case study organizations. In a full strategy, each section would have multiple entries (as suggested in Worksheet 10), but the case studies below provide sample language that may be useful to your organization.

> *We affirm our commitment to equity throughout the lifecycle of our data.*
> *The following are the principles under which we operate with respect to data:*

CASE STUDY A

ORGANIZATION A uses Worksheet 5 to draft the following:

We recognize that data represents people and commit to treating our clients' data with the same respect we extend to their person-hood, including valuing self-identification. One concrete activity that demonstrates this commitment is engaging clients in determining our data standards, such as population group identifiers. Another is recording clients' preferred name and pronouns and using them in any context where they feel comfortable having them used.

CASE STUDY B

ORGANIZATION B uses Worksheet 5 to draft the following:

Our organizational integration of GBA Plus includes a commitment to challenging our assumptions. Because we recognize the role bias can play in data processes, we have made this a priority item for our Equity Committee. We draw on the lived experience of clients through mechanisms like an annual client satisfaction survey to put our quantitative data in context and challenge our assumptions about different groups of newcomers and their experiences with our services.

> *Our agency has identified the following purposes for the data we collect:*

CASE STUDY A

ORGANIZATION A refers to Worksheet 2 to draft the following:

As part of providing the best possible services for our clients, we use data to understand who we are serving in order to build relationships with community organizations who provide similar services or support overlapping populations.

CASE STUDY B

ORGANIZATION B draws on Worksheet 1 to draft the following:

We use data to understand how diverse groups of newcomers experience our services, connecting demographic data collected at client intake with progress notes and satisfaction surveys to see how intersecting identity factors impact service usage and benefits.

To support these practices and principles, we have developed the following policies and guidelines:

EQUITY-FOCUSED POLICIES

ORGANIZATION A brings together the evaluation action steps from Worksheet A2 and the design from Worksheet 4 to draft this section:

One key step to ensure access to quality data is collection. We train our staff to collect data in a culturally safe manner through strong rapport and cultural humility, ensuring staff have a strong understanding of informed consent, including why each piece of data is collected, how it is protected and used, and what it means for a client to consent to sharing their personal data.

CORE PRACTICES

ORGANIZATION B refers to Worksheet 7 to describe this measure:

To ensure high-quality data, we conduct twice-yearly audits of all active client files to ensure our data is complete, accurate, and valid.

Data Practices Evaluation Guide

The following section outlines the core practices of data management.

An evaluation of data practices should address each function across the data lifecycle to ensure that data processes meet all applicable requirements as well as supporting agency data needs and goals (as identified on Worksheet 1 and Worksheet 2). The Data Evaluation Worksheets in Appendix B include guiding questions for agencies to use in their evaluation process, while this section provides more information about each core data practice alongside suggestions for possible activities when carrying out the evaluation. The Key Data Terms may also be helpful when reviewing this content.

Standards

Data standards are the rules used to standardize how data is described, represented, and structured. They include common terminology (such as gender categories), definitions (e.g. what constitutes a disability), representation (for example, dates), and the process of making data recognizable through common file formats. Data should be findable, accessible, interoperable, and (re)usable.

Because agencies may want their internal data to be comparable to data from other sources, such as IRCC data, it is recommended that agencies consult governmental data standards, many of which are available through the Government Open Data portal and on the Statistics Canada website. Organizations in the same sector and/or province or region may also benefit from cross-agency collaboration in establishing shared data standards, which better enables data sharing and comparisons to generate a larger and more accurate picture of the sector as a whole.

WHAT'S COMMON?

Common file formats include .xls and .csv (both of which can be opened using spreadsheet programs like Excel). There are a wide variety of file formats out there! Don't get bogged down with the variety—consider who you share data with and what formats these partners can read.

In addition to government standards, agencies can consult with staff of diverse backgrounds as well as the communities they serve to learn what categories and labels accurately reflect that community's self-identification. In order to ensure data standards are used correctly, agencies should also maintain clear documentation with relevant definitions for staff. Agencies may consider creating similar resources for clients to increase transparency and accountability.

Data Practices Evaluation Guide

POSSIBLE ACTIVITIES:

- Examine the standards currently in use to identify any areas where your data is not standardized.

- Compare existing agency standards to those of entities whose data the agency wants to be able to compare against, such as the Government of Canada or other organizations with whom the agency wants to collaborate.

- Consult with staff of diverse backgrounds and identities as well as client communities to better understand the labels and categories used in self-description.

- Connect current standards, collaborator standards, and self-identification to update data standards for the agency.

- Create a guide to agency data standards and hold trainings for staff who collect, enter, and use data.

Collection

Data collection is a pivotal part of the data lifecycle. Staff collecting data should be trained to understand what they are collecting and why; the nuances of informed consent; how the data will be used, who can see it, where and how long it will be kept; and in establishing culturally safe spaces. Federal and provincial legislation may affect what data can or should be collected. For example, in British Columbia, it is recommended that demographic data be collected, but refusal to answer demographic questions cannot disqualify clients from accessing services. Agencies should review any data-related legislation or guidelines in their respective province or relevant contracts to ensure that they are complying with regulations and stipulations.

An understanding of what data is being collected, and the purpose it serves, is necessary for staff to have and to be able to share with the client. Stronger understanding may improve client receptivity and increase response accuracy, as staff will be better able to explain the meaning of questions or categories which may be unclear to the client. Understanding how the data will be used, stored, and protected is also important information to share with the client, as all of these things are integral to informed consent.

Culturally safe data collection centres on a recognition not just of possible cultural differences, but also the power relationship between staff and clients and the potential for data collection to harm individuals or communities.[6] For many clients, staff are the gatekeepers of services. Newcomers may be afraid of stigma, losing access to needed supports, or even being sent back to a country from which they fled,[7] a concern that is particularly among refugees with disabilities.[8] Understand that the mere possession of sensitive information is a form of power. For example, some newcomers do not share their gender identity or sexual orientation with their co-ethnic communities, fearing a loss of support if this information were to be known.[9] A settlement worker with this data then has the ability to cause substantial harm to that client by sharing this information. It requires care and humility to create a space where clients feel safe sharing this kind of sensitive information.

THE RIGHT TO BE FORGOTTEN

Informed consent is essential to ethical data work. But people may change their minds about sharing their personal information, so you need a way for clients to retract their consent (also called "the right to be forgotten"). This is often a simple written form. Let clients know how this works at the same time as you ask for their consent. Remember to build in a method for removing identifying data so that you can honor these decisions.

POSSIBLE ACTIVITIES:

- Review federal and provincial regulations and contract stipulations related to data collection.

- Examine data collection practices within the agency, including who is collecting data, the environment in which data is collected, and how informed consent is navigated.

- Locate any gaps in institutional knowledge and staff training with respect to culturally safe and trauma-informed data collection.

- Establish guidelines and training on culturally safe and trauma-informed data collection.

Data Practices Evaluation Guide

Quality

Data quality refers to the accuracy, completeness, and consistency of data. High-quality data increases trustworthiness and ensures that evidence-based decisions are made on a strong foundation of evidence. Agencies should establish a clear internal expectation for data quality, measured in percentages. In the initial evaluation, and at regular intervals, data quality should be assessed and protocols adjusted to support high quality data using the following measures.

Accuracy refers to the validity of data. For example, if a data field has an invalid value, it may have a word instead of a number, or it might contain an impossible answer, such as an age of 230. Not all unusual values are necessarily wrong, but it is important to determine the proportion of unusual values and to identify and correct the ones that are definitively wrong, if possible.

Completeness is a measure of how many data fields are filled or empty. Incomplete data impedes analysis, so it is important to know how complete your data is, and to understand how incomplete data may affect any analysis or derived data. Include this measure in your metadata.

Consistency means that all data is using the same standards. For example, all entries use the same options for gender, and each option is defined the same across entries. Inconsistencies reduce comparability and thereby complicate and impede analysis and derivation.

The Government of Canada recommends the VIMO standard, measuring Valid versus Invalid data, Missing and Outlier values. These can be measured using simple data analysis and visualizations, such as generating a scatter plot that would clearly display outlier data.

GRAPH IT OUT

Simple data visualizations (like scatter plots or bar graphs) can be generated through spreadsheet software like Excel. These can be a quick way to check for outlier data—you will see your unusual data point off by itself and can check to see if it is a valid outlier or if it is the result of an error.

POSSIBLE ACTIVITIES:

- Audit data quality throughout organization.
- Identify areas for improvement.
- Establish data quality targets.

Curation

Agencies must consider federal and provincial privacy and protection laws and regulations as applicable to their data practices. This is especially important for managing data privacy and related protections. Because such regulations may change over time, it is essential for agencies to incorporate a review of data and privacy laws into their ongoing monitoring and evaluation practices. The Office of the Privacy Commissioner of Canada maintains a website, https://priv.gc.ca, with an overview of privacy-related laws and guides, including links and resources for privacy laws and oversight in each province and territory. In addition, contracts with federal and provincial governments may include stipulations around data privacy and protections, such as a requirement to store data within Canada. Agencies should review all funding contracts for privacy and protection requirements or guidelines.

Data Practices Evaluation Guide

Warehousing is an important step in the curation process. Evaluating this stage of the data lifecycle requires examining where and how data is stored, both digitally and physically, and for how long it is retained before being deleted or destroyed. As with privacy and protection, agencies should consult federal and provincial regulations as well as relevant contracts. In addition to privacy protections, this step includes determining who, inside or outside the agency, has access to data and in what ways. This encompasses entering, viewing, changing, deleting, and exporting data. It is recommended that access be controlled, and that all agency staff with data access receive training appropriate to their data-related roles, including expectations and procedures for protecting client privacy. As data entry is one point at which inaccuracies may enter, training and quality checks for data entry are recommended.

Database structures play a key role in both curation and analysis. Because GBA Plus integration requires access to disaggregated data, agency databases need to provide mechanisms for disaggregating the data they contain. To avoid manual collation, which is time-consuming and error-prone, it is recommended that agencies establish database reporting structures for the desired disaggregation and collation, allowing the appropriate staff to extract data in the desired combination and format.

WHAT MAKES DATA CLEAN?

Data cleaning is simply the process of correcting or removing data that is inaccurate, incorrectly formatted, incomplete, or corrupted. Cleaning is the next step after checking data quality.

Curation also includes cleaning and processing data for use. Agencies need to have an established method for handling missing and/or inaccurate data. This can be handled in several ways, but it is important to center data equity in determining protocols for managing missing and inaccurate data. Excluding a data record because it is missing one or more fields may perpetuate underrepresentation, while filling in a "best guess" to correct a field may contribute to erasure and propagate biases and assumptions. One option is to follow up with the client in question to verify the answer.

Metadata is an essential part of curation, as it makes data easier to find, interpret, trust, and use. Metadata provides information about the source of the data (such as how and when it was collected), what processing was done (such as how it was cleaned or how aggregate numbers were calculated), an assessment of the data's quality, and the data standards that were followed. This information provides context for analysis and plays a valuable role in protecting against misinterpretation. This is important to include when reporting on data. Check with your database vendor to see what options you have for automatically generating some of your metadata.

POSSIBLE ACTIVITIES:

- Review federal and provincial legislation and funding contracts for stipulations regarding privacy and protection.

- Assess current practices against legal and contractual requirements.

- Determine whether existing database structures support all of the agency's needs and the extent to which they enable disaggregation and GBA Plus analysis.

- Audit curation practices for handling missing, invalid, or inconsistent data. Re-examine these practices through a data equity lens.

- Audit metadata practices for their ability to support finding, interpreting, trusting, and using data.

- Establish clear guidelines and practices for each aspect of data curations.

Data Practices Evaluation Guide

Analysis & Use

For data management to support equity and intersectional analysis, it needs to enable access to disaggregated data, which is pivotal to any equity-based analysis, as it is difficult to find and identify gaps without the ability to compare service use or outcomes across demographic groups. Agencies may want or need to perform many kinds of analysis in the course of their work, and it is essential to retain an intersectional equity-driven lens throughout all analysis. The Government of Canada guiding principles of data ethics include using data to benefit society as well as ensuring that information derived from the data is reported fairly and does no harm.[10]

How information is presented, even the words used to frame it, impacts how that information is received and interpreted. It is essential to center the people from whom data has been collected and to consider how framing and language choices empower or impede the people and communities being described. When partnering with other stakeholders, agencies need to bring these concerns to the collaboration. Equity should drive collaboration, from selecting partners to shaping the relationship and the work done together.

Much like the data itself, analysis does not exist in a vacuum. It is performed for a purpose and put to use, both of which should incorporate an intersectional lens.

- Is this data being used to identify and address equity gaps or to excuse them?
- How could analysis be interpreted or used by other parties to support or hinder equity-seeking groups?
- In what ways can agency data analysis empower newcomer communities?

Reports and analysis cannot be controlled after they have been released, so these questions need to be considered beforehand.

STRIVING OR THRIVING?

It's very easy to fall into the trap of presenting all of your data around deficits by focusing on what is lacking. But that creates a very skewed image of newcomers. Consider ways to share successes, triumphs, joys, and the many ways newcomers thrive and contribute to their communities.

Analysis and utilization of data is a process that benefits greatly from engagement with client communities. Qualitative data from interviews, focus groups, and open-ended surveys provide valuable context, reinforcing or challenging results derived from quantitative data. Quantitative data is an excellent tool for identifying gaps, while qualitative data from community engagement is often required to determine the specific causes and barriers behind these gaps. For example, it is not enough to note that trans women are underrepresented in language classes; it is necessary to follow through on that observation and engage directly with trans women clients to learn why there is lower uptake. Do trans women fear or face discrimination from instructors or fellow students? Do they tend to have more challenges related to basic needs like income, food, and housing that make attendance more difficult? Without direct engagement, efforts to address this service gap may be misdirected and unsuccessful.

POSSIBLE ACTIVITIES:

- Identify the frameworks used to guide data analysis conducted by the agency. Consider the role data equity and GBA Plus play, or could play, in these processes.
- Examine how the agency engages with newcomer communities around data and analysis
- Consider which entities the agency shares its analysis, and the format in which the analysis is presented. Investigate possible mechanisms and accessible formats for sharing analysis with equity-seeking communities.
- Explore ways to engage newcomer communities directly in order to understand how data use impacts them.

Data Practices Evaluation Guide

Monitoring & Evaluation

Finally, agencies must consider existing evaluation and monitoring practices.

- Are there procedures already in place to assess data quality?

- To monitor data collection?

- To maintain compliance with legal and contractual privacy obligations?

- If there are procedures, how consistently are they followed, and at what intervals?

- What practices are in place to identify and correct bias at each stage of the data lifecycle?

This evaluation framework can be a valuable tool for agencies seeking to improve their data practices, but evaluation and monitoring must be ongoing. Data standards may evolve to more accurately represent client identities, or to more closely align with outside standards, such as those used by the Government of Canada. Data collection should be monitored to ensure that it is gathered in a respectful and appropriate manner. Data quality needs to be regularly assessed. Curation practices may require updating to comply with legal and contractual obligations, to respond to evolving technology, or to prepare for disaggregation on alternate axes. Analysis and data use need to be closely monitored to uphold data ethics and agency values. At every stage, these processes need to remain grounded in data equity principles and be structured to facilitate intersectional analysis.

IF YOU LIKE IT, KEEP IT

This guide is intended to be reasonably comprehensive, but its component parts can be pulled apart and adapted in whatever way works for you. Maybe you want to have a yearly staff engagement event where you revisit your data questions and narratives, or evaluate one core practice every six months. If it works for you, do it!

POSSIBLE ACTIVITIES:

- Examine any existing monitoring and evaluation procedures for suitability, thoroughness, and usefulness. Consider where they are meeting agency needs as well as where gaps might arise.

- Consider streamlining relevant portions of this evaluation guide to fit agency needs and capacity as a regular activity.

Data Practices Evaluation Worksheets

The worksheets in this section are designed to provide guiding questions and possible activities for organizations as they evaluate their core data practices (as described in Phase Two and elaborated upon in Appendix A. Resources and capacity may limit the extent of an evaluation, so use as many or as few of these worksheets as makes sense for your data journey. Keep Worksheet 3 prominent and visible throughout this process, to sustain your reflection on equity and values across data practices.

Worksheet A1: Evaluating Data Standards

Data Standards Questions:	Notes:
Data Standards Questions: • What terms, categories, and variables are in use? (e.g. gender identities, ethnocultural identifiers, etc.) • Are all client identities represented in these standards? • What mechanisms are in place to update terms? • How are data standards communicated to staff? • How consistently are these standards used and upheld? • How are these standards expressed in the metadata? **Possible activities:** • Refer to data standards used by federal and provincial governments • Consider consultation with ethnocultural groups to determine terms. • Avoid using "other" as a catch-all category. • Provide clear documentation for all data standards in plain language.	

Data Practices Evaluation Worksheets

Worksheet A2: Evaluating Data Collection

Data Collection Questions:

- Who is collecting data?
- Does the staff composition reflect the diversity of the population served?
- What are your data collection guidelines?
- How do they reflect agency values and data needs?
- What training do staff have? What training is the agency currently offering?
- What data is being collected? From whom?
- What are possible harms from this collection?
- How are those harms mitigated?

Possible activities:

- Consider updating data collection guidelines or practices to reflect emerging priorities.
- Consider offering additional training in data collection and cultural safety.
- Ensure all data being collected is minimally invasive and necessary for answering priority questions.
- Consider consulting with clients and communities about possible harms and concerns.

Notes:

Data Practices Evaluation Worksheets

Worksheet A3: Evaluating Data Quality

Data Quality Questions:

- What is the expected standard of data quality?

- How is accuracy checked?

- How accurate is the current data?

- How is data completeness checked?

- How complete is the current data?

- How is data consistency checked?

- How consistent is the current data?

- How is missing or inaccurate data handled?

Possible activities:

- Determine what standard of data quality is desired and what is acceptable.

- Develop a standard quality check procedure/audit process and institute a regular schedule.

- Develop agency guidelines for handling missing and/or invalid data.

Notes:

Data Practices Evaluation Worksheets

Worksheet A4: Evaluating Data Curation

Data Curation Questions:

- What legislation or contract stipulations exist concerning data storage and privacy? Where is data stored? What protections are in place?

- Who can access, change, and pull data?

- How is anonymization handled?

- Can data be disaggregated to the degree necessary to answer priority questions? How can future disaggregation be introduced?

- How, with whom, and for what purposes is agency data shared? What processes are in place to produce shareable data?

Possible activities:

- Refer to data storage and protection guidelines outlined in federal and provincial legislation and funder contracts.

- Implement a regular check to ensure protections meet standards and are up to date.

- Develop practices to ensure adequate disaggregation and a procedure for implementing new disaggregation protocols as necessary.

- Determine with whom the agency needs or wants to share data.

- Establish guidelines to ensure data can be exported in formats usable by data partners.

Notes:

Data Practices Evaluation Worksheets

Worksheet A5: Evaluating Data Use

Data Use Questions:

- What methods are used to analyze data?

- How is qualitative data integrated into data analysis?

- What processes are in place to ensure data is reported fairly and used responsibly?

- How is data currently integrated into agency decision-making?

- What partnerships are in place, and what role does data play in them?

- How are client communities connected to the ways agency data is used?

Possible activities:

- Establish guidelines for contextualizing data and using qualitative data (e.g. focus groups, open-ended surveys) to check agency data against client experiences.

- Integrate data equity oversight into reporting.

- Consider agency decision-making models in the context of data equity.

- Embed GBA Plus and data equity principles in partnerships and collaborations.

- Consider community outreach and engagement practices.

Notes:

Worksheet A6: Evaluating Data Monitoring

Data Monitoring Questions:

- What processes are in place to monitor data quality? What processes are in place to assess practices for potential harms?

- How is the agency accountable to its clients and funders?

- How diverse is the team responsible for data throughout its lifecycle? What measures are in place to ensure diverse perspectives?

- How often are data security measures evaluated? In what ways is quantitative data checked against qualitative data?

- What resources are available to staff for ongoing skills development?

Possible activities:

- Update or implement processes and schedule for regular data quality monitoring and updating data security measures.

- Update or implement standard practices for community consultation.

- Integrate accountability standards and measures into monitoring and evaluation framework.

- Consider regular qualitative data collection and comparison between qualitative and quantitative data for gaps and incongruencies.

- Consider developing, or partnering to create, resources and training for staff involved in the data lifecycle.

Notes:

References

Bopp, Chris, Ellie Harmon, Amy Voida (2017). Disempowered by Data: Nonprofits, Social Enterprises, and the Consequences of Data-Driven Work. Chi '17: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems.

Das, Meena (n.d.) Ten community-centric data principles. Namaste Data.

D'Ignazio, Catherine and Lauren Klein (2020). What gets counted counts. *Data Feminism.* The MIT Press. Retrieved February 29, 2024, from https://data-feminism.mitpress.mit.edu/pub/h1w0nbqp/release/3

Janzen, Rich and Joanna Ochocka (2020). Assessing excellence in community-based research: Lessons from research with Syrian refugee newcomers. *Gateways: International Journal of Community Research and Engagement 13*(1).

Kahn, Sarilee and Edward J. Alessi (2018). Coming Out Under the Gun: Exploring the psychological dimensions of seeking refugee status for LGBT claimants in Canada. *Journal of Refugee Studies 31*(1):22-41.

Komtheuer, Annette, Michaela Hynie, Martha Kleist, Safwathullah Farooqui, Eva Lutter, and Manuela Westphal (2021). Inclusive resettlement? Integration pathways of resettled refugees with disabilities in Germany and Canada. *Frontiers in Human Dynamics 3.*

Lee, Edward & Shari Brotman (2013). SPEAK OUT! Structural intersectionality and anti-oppressive practice with LGBTQ refugees in Canada. *Canadian social work review/Revue Canadienne de service social 30*:157-183.

Nickerson, Marcia (2019). Cultural safety and humility: Case study report. First Nations Health Authority.

Endnotes

1 Modules on data literacy, including data ethics, are available from Statistics Canada in English at https://www.statcan.gc.ca/en/wtc/data-literacy and in French at https://www.statcan.gc.ca/fr/afc/litteratie-donnees

2 Das, Meena (n.d.) Ten community-centric data principles. Namaste Data.

3 D'Ignazio, Catherine and Lauren Klein (2020). What gets counted counts. *Data Feminism.* The MIT Press.

4 Das, Meena (n.d.) Ten community-centric data principles. Namaste Data.

5 Bopp, Chris, Ellie Harmon, Amy Voida (2017). Disempowered by Data: Nonprofits, Social Enterprises, and the Consequences of Data-Driven Work. Chi '17: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems.

6 Nickerson, Marcia (2019). Cultural safety and humility: Case study report. First Nations Health Authority.

7 Janzen, Rich and Joanna Ochocka (2020). Assessing excellence in community-based research: Lessons from research with Syrian refugee newcomers. *Gateways: International Journal of Community Research and Engagement 13*(1).

8 Komtheuer, Annette, Michaela Hynie, Martha Kleist, Safwathullah Farooqui, Eva Lutter, and Manuela Westphal (2021). Inclusive resettlement? Integration pathways of resettled refugees with disabilities in Germany and Canada. *Frontiers in Human Dynamics 3.*

9 Lee, Edward & Shari Brotman (2013). SPEAK OUT! Structural intersectionality and anti-oppressive practice with LGBTQ refugees in Canada. *Canadian social work review/Revue Canadienne de service social 30*:157-183; Kahn, Sarilee and Edward J. Alessi (2018). Coming Out Under the Gun: Exploring the psychological dimensions of seeking refugee status for LGBT claimants in Canada. *Journal of Refugee Studies 31*(1):22-41.

10 Statistics Canada (2022). Data ethics: An introduction. Retrieved February 29, 2024, from https://www.statcan.gc.ca/en/wtc/data-literacy/catalogue/892000062022001

AMSSA®
Strengthening Diversity In BC

AMSSA@AMSSA.ORG | AMSSA.ORG