Technical Brief 4: Connecting Gender Data to Action

Authoured by Alex Berryhill and Lorena Fuentes, Ladysmith (April 2022)

This brief was developed in partnership with Gender at Work for the Global South AI4COVID Program which is funded by The International Development Research Centre (IDRC) and the Swedish International Cooperation Agency (Sida)

Recent international initiatives[1] show an increasing concern with gender data gaps, and—thanks in large part to decades of feminist advocacy—a growing awareness of the significant harm that can be caused by research that excludes women, girls, and other marginalized communities' experiences.[2] These gaps have become especially evident in the wake of the COVID-19 pandemic.[3] While this momentum to address global gender data gaps is highly welcomed**, today's data ecosystem remains largely built around a false "build it, and they will come" assumption**[4]—with greater investments being made in data collection and publication, rather than gender data uptake and impact. Indeed, a recent study by Data2X indicates that organizations involved in increasing the uptake and impact of available data make up the smallest share of stakeholders in the global gender data industry.[5]

This is a concerning trend, as **gender data alone will not catalyze gender equality: data must be connected to action that puts power into the hands of women and girls**.[6] Having more and better gender data does not guarantee that the data will actually be used to inform decisions, nor that these decisions will lead to positive gender equality outcomes. For example, although national statistics offices and institutions are increasingly leveraging time use surveys to collect data on unpaid care work, social protection systems continue to overlook the needs of paid and unpaid care workers—as has been made especially evident in the wake of the COVID-19 pandemic.[7]

Connecting data to action requires moving far beyond data collection and publication: **investment is needed to facilitate the uptake and impact of gender data as well.** For gender-transformative data projects, this requires developing and following a clear Theory of Change, grounded in the local context: What are the project's intended outcomes? What activities, beyond data collection, are needed to realize these outcomes? Through this process, gender data projects will often need to identify and address more than just data gaps, but also related and intersecting gaps in political will, technical capacities, resource allocation, and empathy for the issues and communities under

[1] See Cookson & Fuentes 2021 for examples of recent international investment in resolving global gender data gaps.
[2] Criado-Perez 2019; D'Ignazio & Klein 2020; Fuentes & Cookson 2020
[3] For example, see UN Women 2020; ILO 2020; Data2X 2020; and McDougal et al 2021.
[4] Open Data Watch & Data 2X 2019; Stern 2017.
[5] Data 2X 2021.
[6] Cookson & Fuentes 2021
[7] Data 2X 2018; ILO 2019; UN Women 2020

consideration.[8] With the state of today's gender data ecosystem in mind, this brief provides guidance on how to strategically invest in data uptake and impact, in order to help research projects contribute to and achieve positive gender equality outcomes.

How are the partners of the Global South AI4COVID program connecting gender data to action?

The **Argentinian Public Health Research on Data Science and Artificial Intelligence for Epidemic Prevention (ARPHAI)** initiative develops data-driven tools for informing public health policy, including state responses to current and potential future pan/epidemics. These tools largely draw upon aggregated and processed electronic health record (EHR) data. Through ARPHAI´s processing and analysis methodologies, they leverage EHR data to respond to health crises - while protecting such highly sensitive data and ensuring its fair representation.

Policy impact has been a driving force throughout the design and implementation of the ARPHAI initiative. Indeed, ARPHAI has invested in multiple strategies to connect their data to action:

- **Identifying and responding to data users' capacity gaps:** ARPHAI identified potential barriers to leveraging health data through discussions with their multidisciplinary team–which includes investigators engaged in international clinical research (a highly regulated field, with generally more established and stricter data protection standards than international AI research). Through dialogue across different disciplines, they found that one of the key barriers to using health data for more evidence-based and inclusive policymaking is the lack of capacity around *safely* using such highly sensitive data. Responding to this barrier, ARPHAI reached out to public sector stakeholders (ranging from project managers to data analysts, and individuals with high school degrees to researchers with doctorates), to offer workshops on privacy and ethics issues in health data management. However, ARPHAI soon identified another barrier: stakeholders' time constraints. Responding to this reality, ARPHAI is now working on **action-oriented guidelines and resources** on responsible health data management, designed specifically to meet stakeholders' needs. From data intake to report generation, these guidelines outline methodologies and examples for the secure and fair handling of health data. Once completed, these guidelines will be made available on the Global South AI4COVID website.

- **Facilitating communication between data generators and users:** Having members of the ARPHAI team on-site, located at the institutions where data are originally generated (e.g Primary Healthcare Centers implementing the EHR systems under analysis) has also been key to allow for a fluid communication between those who generated the data and those who analyze and use it. Likewise, having a **multidisciplinary team that can communicate**

[8] Fuentes 2020

effectively with data-generating personnel, engineers, system administrators, public health officials, medical doctors, administration managers, and researchers from different scientific fields has also helped facilitate ARPHAI's communication with data users, and thus ARPHAI's ability to connect their research to positive policy outcomes.

- **Identifying and responding to strategic 'entry points' for change:** Lastly, as discussed in Technical Brief #3, ARPHAI's close engagement with policymakers has allowed their team to identify and respond to political contexts and opportunities for impact. For example, after analyzing the EHR system currently under development by the Ministry of Health, and **engaging with policymakers** from the Ministry of Health and the Ministry of Women, Gender and Diversity in Argentina, the ARPHAI team identified the need for a new data model that integrates nonconforming gender identity categories. ARPHAI responded to this need by carrying on research to generate evidence-based recommendations supported by previous experiences, guidelines from other countries, and the views from different experts and policy makers on that matter. This work is currently being considered by the Ministry of Health to improve nonconforming gender identity representation in their EHR system—illustrating a significant advancement towards fulfilling the aspirations of the country's 2012 Gender Identity Law.

Key challenges and solutions to connecting gender data to action:

1. **Decision-makers lack the skills needed to interpret and understand gender data.** Studies show that data fluency continues to be a key barrier to data uptake for a diverse range of stakeholders.[9] Formal (such as policymakers) and informal (such as community-based leaders) decision-makers may not always be equipped with the time, resources, or skills necessary to interpret gender data in a way that will allow for them to then act upon this data in a useful way.

   → *To overcome this challenge: This is why engaging with formal and informal decision-makers from the beginning of a project's design is key. When projects are designed alongside* ***intended data users*** *and* ***gender equality and women's rights stakeholders*** *from the very beginning, there is greater opportunity to shape data collection, analysis, as well as publication and dissemination around their unique capacities.*[10] *For example, gender equality and women's rights stakeholders are well positioned to have a deeper understanding of the local political context, and can therefore help identify what type of data (and type of data*

[9] Elrha 2021; Open Data Watch & Data 2X 2019; Nesbitt 2021; Custer & Sethi 2017.

[10] For more on the importance of engaging with gender equality and women's rights stakeholders for designing data uptake plans, see Cookson & Fuentes 2021 and Zulver et al 2021.

communications) are most needed to advance gender equality. Close engagement with relevant stakeholders helps teams design data projects that are more relevant and responsive to stakeholders needs—while recognizing that these needs may change throughout a project's lifecycle.

→ Additionally, project plans should allow ample time and resources for ***strategic data communication****. Data may need to be presented in different ways for different audiences, depending on different capacities or expectations for data use. As illustrated by ARPHAI's experience (see box above) having a multidisciplinary team that can communicate effectively with data producers and users is also key for connecting data to action.*

2. **Decision-makers lack the political will to respond to data findings.** Often, where gender data gaps are present, so too are gaps in political will and/or empathy for the issues and communities under consideration. Evidence indicates that when data collection is driven by external funding, it is less likely to create impact compared to data collection supported by local decision-makers and their respective institutions.[11] This is likely because supporting data collection on a particular issue indicates that there is already local buy-in and political will to better understand this issue. Stakeholders' likelihood of responding to or using gender data findings will also be dependent on local dynamics: Does their institution support or provide incentives for evidence-based decision-making that advances gender equality?

   → To overcome this challenge: During a project's design phase, a ***stakeholder mapping exercise*** *can help identify gaps in political will, as well as opportunities for strategic partnerships that may help establish or amplify political will, including by surfacing joint agendas between stakeholders with seemingly different priorities and incentives (e.g., grassroots women's organizations and municipal/local government).[12] Based on this mapping, projects can design research strategies that build upon and strengthen political will where it exists—and where it doesn't, they can consider other strategies that may help establish that political will.*

How are the partners of the Global South AI4COVID program connecting gender data to action?
The **Université Cheikh Anta DIOP's (UCAD)** research seeks to support epidemiological modeling of COVID-19 in Senegal and Mali, while also identifying and addressing the social and political challenges of using AI technologies. Part of this research has included a mixed-method survey

[11] Data 2X 2018; Custer & Sethi 2017.

[12] For more on stakeholder engagement, see Technical Brief #3. For guidance on stakeholder mapping specifically, see Lusthaus et al 1999.

distributed periodically in Senegal throughout the year (2021). Each round generated 1,000 survey responses, which were coupled with 200 in-depth interviews with key stakeholders. By **collecting a mix of qualitative and qualitative data**, the UCAD team has been able to uniquely capture the pandemic's differentiated impacts based on gender, age, occupation, and geography, among other factors—as well as potential drivers of these differentiated impacts. For example, their research has found that the state's social protection programs have largely excluded the country's informal workers. Given women's overrepresentation in the country's informal sector, they were less likely to have access to these benefits, and thus faced even greater vulnerability to the pandemic's economic impacts.

While the team is currently in the data analysis phase of their project, they are also thinking ahead to how they can ensure data uptake so that their gender data may inform national policymaking. To do so, they have identified a **strategic entry point for strengthening national political will** around more gender equitable pandemic response policies: Gender specialists within relevant national ministries, such as the Ministry of Women, Family and Children, Ministry of Health and Social Action and from women's rights and child protection civil society organizations. Through **interviews with key stakeholders and analysis of the political context**, UCAD's team has identified these gender specialists across relevant institutions as a key pathway for their gender data to eventually inform policymaking. Ministry-based gender specialists often lack the data they need to advocate for a more gender equitable response, and UCAD's data can help address this data gap. As such, UCAD plans to **develop strategic data communication products targeting these stakeholders.** Webinars and policy briefs will be designed bearing in mind the specific needs and capacities of ministry and civil society-based gender specialists, in order to support the effective uptake of their findings.

3. **Potential risks from data collection and dissemination.** When considering what data is needed to inform decisions and support actions for positive gender equality outcomes, project designers and researchers must also consider the benefits and risks to data collection. More data on marginalized communities may help visibilize their lived experiences and thus support actions to better meet their needs. However, greater visibility may also create new risks for already vulnerable populations. For example, while collecting data on sexual orientation and gender identity (SOGI) can help visibilize disparities based on SOGI and thus support actions to respond to these disparities, visibility without guarantees of safety and protection can catalyze new risks or create additional stigmas against LGBQT+ communities.

   → To overcome this challenge: ***Data minimization*** *is key to reduce protection risks. Researchers must first ask alongside data subjects and impacted communities*: *What data is needed to support needed actions? Then*: *What are the risks associated with this data*

collection? Are there opportunities to mitigate or minimize these risks? Impacted communities must be engaged across the process. Impacted communities should also be involved in monitoring the outcomes of these decisions: Have new risks been identified during data collection, analysis, or use? Do these new risks alter the project's ***risk-benefit analysis****?*

→ Additionally, supporting data users' awareness of data protection concerns, as well as ***building capacity around rights-based data management*** *can also help decrease potential risks around data use. Indeed, ARPHAI's experience indicates that such capacity-building is key for facilitating effective research uptake, especially when working with sensitive data.*

Descriptive vs. explanatory data: Matching gender data to desired outcomes

Not all data is equally 'actionable'. Different types of data will allow for, or may help inform, different types of actions. In order to identify what type of data is needed to advance gender equality goals, it is important to distinguish between descriptive and explanatory data—which each serve distinct purposes.

> ***Descriptive data*** is useful for **monitoring progress towards gender equality**, making visible gender disparities, and 'making the case' for gender equality investments. Descriptive data may be qualitative or quantitative. However, descriptive data that *lacks* qualitative research methods risks overlooking important local dynamics to observed statistical trends.[13] For example, sex-disaggreated statistics on school attendance may illustrate gender disparities in school access, while interviews with stakeholders may identify some of the negative consequences fo these gender disparities. Both of these descriptive data sources can then be used to raise awareness of gender gaps in education, and advocate for state investment to decrease identified gaps between boys and girls' access to education.

However, while descriptive data helps visibilize problems, it is insufficient for designing *solutions* to these problems. In order to design context-appropriate policies and programs, we also need explanatory data.

> ***Explanatory data*** helps us analyze the **root causes and drivers of gender inequality**, and in doing so, identify strategic opportunities for action. For example, returning to the example above, explanatory data might include key informant interviews or participatory workshops with students, parents, and teachers (disaggregated by sex). Such data sources would help to capture *why* there are gender disparities in school access in a particular local context and *where* there are potential, context-appropriate entry points for change.

No single data source or research methodology is sufficient to capture "the richness of insights required to advance gender equality"[14]—nor is any single data source sufficient for driving action.

[13] Cookson 2018; UNICEF 2020.
[14] Haan et al 2020.

We need descriptive data to understand local gender (in)equality dynamics, mobilize gender equality investments, and track progress towards gender equality. We need explanatory data—evidence of why and how gender inequalities manifest in unique contexts—in order to then design programs or advocate for specific policy solutions that respond to these inequalities.[15] **Identifying *what* type of data is needed to drive positive gender equality outcomes—as well as the limitations and opportunities of different data sources—is key for successfully connecting gender data to action.**

Key recommendations:

1. **Develop (and then share) a clear Theory of Change for how research will be leveraged for positive gender equality outcomes.** Having a clear ToC helps projects and partners (including local stakeholders and funders) better understand—and thus successfully follow, monitor, and adjust as needed—the project's roadmap for connecting data to action. ToCs are also helpful for unpacking assumptions about data uptake and use, which is key for identifying and responding to potential barriers. Through this exercise, your team will need to discuss and outline: (a) a clear problem-statement; (b) desired outcomes; and (c) outputs and (d) activities needed to realize these outcomes. ToCs should also include evidence that these activities will likely produce intended results; and unpack underlying assumptions. Lastly, engaging data users and local gender equality stakeholders through participatory methods is critical for ground-truthing and strengthening ToCs. This is also a key strategy for developing and strengthening local partnerships from the beginning of a project's design. **For more on how to design a strong Theory of Change for your data project, see the box below.*

2. **From the beginning of a data project's design, establish Memorandums of Understanding (MOU) with local gender equality and women's rights stakeholders.** As discussed in greater depth in Technical Brief #3, engaging decision-makers and power-brokers is key for ensuring data's usability and actionability. Partnerships with both data subjects and intended data users' can help researchers identify data needs, capacities, and potential risks, and then curate project design around these factors (including, for example, developing and vetting the project's proposed ToC). However, establishing partnerships is not sufficient: projects must also facilitate opportunities for stakeholders to meaningfully contribute to research plans and data use. Indeed, IDRC's AI4COVID partners have reported that MOUs have helped formalize spaces to periodically share data findings with stakeholders, thus creating a more enabling environment for partners to share suggestions on how to use data findings, or how to adjust work plans to augment data's actionability. Lastly, engaging specifically with gender equality and women's rights stakeholders helps ensure gender data projects align with and support the

[15] CARE 2011.

needs and priorities of local women's rights movements, which have a greater understanding of and experience with advancing positive gender equality outcomes.

3. **Identify the unique opportunities and constraints to data uptake in that particular context—then develop a plan to respond to these constraints.** Bridging the gap between gender data and action requires attending to the possible barriers and constraints of data uptake in a particular context—and, once these are identified, working as a team to identify possible entry points for data uptake, researcher engagement, or policy influence. For example, in contexts where there is deeply rooted ideological resistance to the concept of gender (or indeed, intersectionality) and therefore risk of backlash, researchers can respond to these constraints and strategically create an entry point for impact by framing gender data in the language of "sex disaggregated data", thus increasing researchers' opportunity to secure meetings wtih key stakeholders. Returning to recommendation #2: Engaging with local stakeholders—particularly intended data users, data subjects, gender equality and women's rights organizations, as well as representatives of potentially impacted communities—is also key for identifying these opportunities, constraints, and potential entry points for impact.

4. **Invest in (timely) data communication.** Delayed publication timelines in peer-reviewed academic journals can present another barrier to connecting data to uptake and impact. Indeed, if projects wait until publication in academic journals, their findings may fail to reach decision-makers with enough time to allow for impact, which is of particular concern in contexts of health or other crises. Instead, projects should consider ways to *also* publish findings in timely strategic communications, such as working papers, policy briefs, or short explanatory videos. This may require investing in a data communications team with expertise in data visualizations, and working with intended data users to ensure presentation of data findings can be easily understood.

5. **Ensure flexibility for iteration, as political contexts change.** Connecting gender data to action requires a deep understanding of local political contexts, which will allow teams to identify and respond to opportunities (and barriers) as they arise. Entry points for change might appear when there are "critical junctures", such as a change of government, or even a change of a single staff member within the Ministry of Health, or a charismatic and compelling gender equality stakeholder who has managed to "catch the ear" of a government decision-maker. However, these junctures are dynamic: it is very likely that the political context from when the project was designed will have changed significantly by the time data is collected and analyzed. As such, projects will need to include flexibility in their timelines, in order to have the opportunity to strategically respond to changing contexts.

Developing a strong Theory of Change for gender-transformative data projects

During a funding or partnership application, donors will often ask for a Theory of Change (ToC), or 'pathway to impact'. However, even if a donor does not specifically ask your team to outline how your project's data will be connected to positive gender equality outcomes, **developing this roadmap alongside your team and project partners will help provide needed clarity on the specific steps needed to meaningfully connect data to positive gender equality outcomes.** This exercise is most effective when completed with project teams, and perhaps even project partners, in an interactive way, with participants encouraged to question assumptions and explore potentially new project possibilities.

The following steps build upon the ToC outline provided by Innovations for Poverty Action (see complete resource here):

1. **Define the problem.** What is the problem your research project seeks to address? Why is your research and data needed? What types of 'gaps' will your project attend to (eg, gaps in data, political will, capacities, resource allocation, and/or empathy)? Back your problem-statement with evidence from a literature review and/or field research.

2. **Define intended outcomes.** What results does your research project intend to achieve? These should be well-defined and specific. For gender-transformative data projects, outcomes should speak to local inequalities and gender norms.

3. **Identify program outputs.** What will your research project need to 'produce' in order to achieve these intended outcomes? Consulting both your team and local stakeholders here will help connect the theoretical with what may be reasonable and feasible in a particular context. Likewise, unpacking assumptions will also help connect what your project can control (outputs) with the changes you intend for them to produce that are outside of the project's control (outcomes).

4. **Define program activities.** What specific activities will need to take place in order to produce these outcomes? Who will be responsible for these activities? Does your team have the necessary skill and resource capacity to realize these activities?

5. **Map pathways between components**. How are the project's activities, outputs, and outcomes connected? Where do multiple components intersect?

6. **Identify assumptions underlying the model.** What assumptions must hold for the research project to work as expected? Returning to the mapped pathways, each 'link' between individual components will likely feature its own underlying assumptions.

Lastly, **a ToC is only as useful as it is understood, monitored, and iterated upon as needed.** Therefore, all project team members and partners should demonstrate a strong understanding of the ToC. Even if a research project does not have a specific monitoring and evaluation team, a team should be established to regularly revisit the ToC, monitor progress, discuss if and how assumptions have changed, and if so—how the ToC may need to adapt.

Additional tools and resources

- Guiding Your Program to Build a Theory of Change (IPA 2016)
- ESPA guide to working with Theory of Change for research projects (Vogel 2012)
- The Data Value Chain: Moving from production to impact (Open Data Watch & Data 2X 2019)
- From Knowing to Doing: Evidence use in the humanitarian sector (Elhra 2021)
- Avoiding Data Graveyards: Insights from data producers & users in three countries (Custer & Sethi 2017)
- How to Engage Stakeholders in Research: Design principles to support improvement (Boaz et al. 2018)