afi

FINANCIAL INCLUSION DATA AND IMPACT (FIDI) WORKING GROUP

SME FINANCE (SMEF) WORKING GROUP

GENDER-RESPONSIVE TOOLKIT FOR MSME DATA ECOSYSTEMS

TOOLKIT

CONTENTS

ACKNOWLEDGMENTS

This toolkit is a product of the SME Finance Working Group (SMEFWG), Financial Inclusion Data and Impact Working Group (FIDIWG), and the Gender Inclusive Finance workstream.

CONTRIBUTORS:

AFI members: Dr Amal Idrissi and the team of the Moroccan SME's Observatory (OMTPME)/Bank Al-Maghrib, who served as the primary methodological and practical backbone of this Toolkit. In addition, the Toolkit is shaped by inputs from more than 50 institutions across the AFI network through surveys and key informant interviews (KII).

AFI Management Unit: Developed under the overall guidance of Nik Kamarun, Senior Policy Manager, SMEFWG, and Luis Treviño, Senior Policy Manager, FIDIWG, whose leadership and coordination were instrumental in the production of this toolkit. The toolkit has gone through gender-lens review contributed by Audrey Hove, Head, Policy Management and Lucy Kabethi, Policy Analyst, Gender Inclusive Finance.

We would like to extend special thanks to **Samuel Hall**, a research organization, for their technical support and contributions to this toolkit.

We would like to thank AFI member institutions, partners and donors for generously contributing to the development of this publication.

The Gender Inclusive Finance workstream is partially inanced by Sweden and other partners.

Cover picture: @ iwciagr / Shutterstock

EXECUTIVE SUMMARY

Women-owned and women-led MSMEs (W-MSMEs) represent an estimated 38 percent of MSMEs globally, yet they face persistent financing gaps, partly because most countries lack robust gender-disaggregated MSME data.

Strengthening these systems enables evidence-based policy, inclusive finance, and improved risk oversight. This toolkit helps AFI members strengthen gender-responsive MSME data systems by providing practical guidance, assessment tools, and good practices. It supports policymakers, regulators, and data producers seeking to integrate gender considerations into MSME definitions, indicators, and reporting frameworks.

TOOLKIT STRUCTURE

1 INTRODUCTION

CHAPTER 1 introduces AFI's approach and alignment with key frameworks.

2 RATIONALE - FRAMING THE CHALLENGE

CHAPTER 2 defines MSMEs and W-MSMEs, highlighting why clear, inclusive classifications are foundational for policy and data comparability.

3 BUILDING BLOCKS OF A GENDER-RESPONSIVE MSME DATA ECOSYSTEM

CHAPTER 3 presents four building blocks of a gender-responsive MSME data ecosystem: institutional foundations, legal and policy enablers, infrastructure and platforms, and data-use pathways.

4 DATA DEVELOPMENT INDICATOR (DDI). ASSESSING THE MSME DATA ECOSYSTEM

CHAPTER 4 applies the Data Development Indicator (DDI), AFI's diagnostic tool for assessing institutional data maturity, and showcases practical solutions.

5 INDICATOR FRAMEWORK AND DATA ARCHITECTURE

CHAPTER 5 proposes a tiered indicator framework (core and enriched) and a prioritization matrix to help members select feasible, high-value indicators.

6 INTERPRETATION OF DATA SYSTEMS AND INDICATORS

CHAPTER 6 provides the basic guidelines for visualizing Data Development Index (DDI) or data ecosystem pathways, and key MSME data indicators.

7 CAPACITY BUILDING AND INSTITUTIONAL STRENGHTENING

CHAPTER 7 offers practical guidance on how institutions can diagnose their data capacity needs by Data Development Index maturity level and dimensions, follow tailored learning pathways, and embed the toolkit approach into their own national data strategies,, working groups, and technical assistance programs.

ANNEXES

ANNEXES provide detailed methodologies, indicator definitions, and resource tools.

INSIGHTS

1 **INSTITUTIONAL LEADERSHIP MATTERS**
Mandates and coordination bodies anchored in national strategies sustain progress.

2 **INCLUSIVE DEFINITIONS DRIVE VISIBILITY**
Consistent criteria for W-MSMEs improve data quality and access to finance.

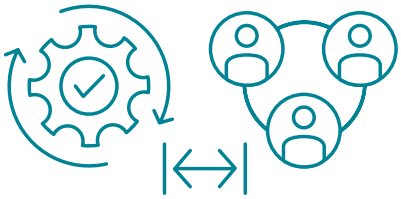
3 **EARLY, LOW-COST WINS BUILD MOMENTUM**
Adding gender fields to registration or surveys creates quick policy value.

4 **TECHNOLOGY BRIDGES GAPS**
Interoperable platforms and AI-enabled tagging expand coverage of informal or legacy data.

5 **PEER LEARNING ACCELERATES CHANGE**
AFI members can adapt lessons from fellow institutions to their own contexts.

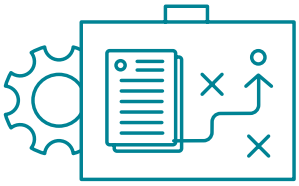
Using this toolkit, members can

Diagnose their data maturity using the Data Development Index



Identify capacity and coordination gaps

Select priority indicators from the tiered framework



Draw lessons from case studies to guide reforms or pilot projects

1 INTRODUCTION

Despite the significant economic role of micro, small, and medium enterprises (MSMEs), W-MSMEs remain disproportionately underserved by financial systems.

38% They represent an estimated 38 percent of global MSMEs[1] yet face a credit gap of USD1.4-1.7 trillion.[2]

This persistent financing gap undermines economic growth, financial inclusion, and gender equity goals, limiting the potential of women entrepreneurs to contribute fully to their economies.

One key driver of this gap is the lack of **gender-responsive MSME data**. In many countries, women-led MSMEs are not systematically tracked, particularly in the informal sector, leaving policymakers and financial institutions without the evidence needed to design targeted interventions. Without reliable, disaggregated data, policymaking risks being gender-blind, and support services may fail to reach those who need them most. Additionally, the absence of such data limits the ability of supervisors and regulators to design differentiated prudential requirements and assess the exposure of financial institutions to vulnerable segments of the MSME market.

The **Alliance for Financial Inclusion (AFI)** has taken a leadership role in addressing this challenge, leveraging its global network of member institutions to advance gender-inclusive finance. The **SME Finance Working Group (SMEFWG)** and the **Financial Inclusion Data and Impact Working Group (FIDIWG)**, both composed of regulators, policymakers, and technical experts, have long championed the need for stronger MSME data systems and provided valuable insights into what works in practice.

It is also important to note that the persistence of this financing gap not only limits economic growth and financial inclusion but may also create blind spots for prudential management. Limited gender-disaggregated information reduces the ability of supervisors and regulators to assess portfolio diversification and exposure to underserved segments.

Strengthening gender data systems can therefore enhance both inclusion and risk oversight within the financial sector.[3]

Purpose and scope of the toolkit

The toolkit is a **practical, modular guide** for policymakers, regulators, and data stakeholders to strengthen gender-responsive MSME data systems at all levels of maturity, recognizing that progress is incremental and shaped by legal, technical, and institutional realities. Rather than prescribing a one-size-fits-all model, it offers adaptable pathways, tools, and examples that can be applied to diverse national contexts.

[1] International Finance Corporation (IFC). 2017. MSME Finance Gap: Assessment of the Shortfalls and Opportunities in Financing Micro, Small and Medium Enterprises in Emerging Markets. Available at: https://documents1.worldbank.org/curated/en/653831510568517947/pdf/121264-WP-PUBLIC-MSMEReportFINAL.pdf

[2] SME Finance Forum. 2019. Women SME Finance Data. Available at: https://www.smefinanceforum.org/

[3] See International Monetary Fund's Gender and Finance: Unlocking Financial Resilience (2022), which discusses how gender gaps in access to finance can limit portfolio diversification and reduce financial system resilience; and also CGAP's Gender Data and Financial Regulation (2021), which highlights how sex-disaggregated data improves supervisory understanding of credit risk and market exposure.

Toolkit target users and use cases

AFI members and partners can use this toolkit in the following ways:

> **POLICYMAKERS AND REGULATORS** seeking to design evidence-based, gender-responsive MSME policies.

> **DATA PRODUCERS** such as national statistical offices, central bank data teams, and SME agencies tasked with collecting and managing MSME data.

> **GENDER SPECIALISTS** working within financial authorities, along with **RESEARCHERS AND TECHNICAL EXPERTS** from institutions that promote financial inclusion.

Use cases include:

- Conducting a **baseline assessment** of an institution's MSME data maturity using the Data Development Index (DDI).
- Identifying **capacity gaps** and designing targeted capacity-building plans.
- Embedding gender-responsive MSME data practices into **national financial inclusion strategies** and monitoring frameworks.
- Supporting **regional and south-south knowledge exchange** on MSME data systems.

Alignment with AFI policy frameworks and global goals

The toolkit complements the AFI's Denarau Action Plan on gender-inclusive finance, supports the AFI Gender Inclusive Finance Policy Model, and aligns with the SME Finance Guideline Note as well as AFI's Policy Framework on MSME Data Collection: A Guide for Gender Inclusive Finance (2023). It further draws on AFI's Data Collection Processes and Defining Micro, Small and Medium Enterprises (2022) and Guideline Note on Sex-Disaggregated Data Report Templates (2021), ensuring consistency with existing guidance on MSME definitions, reporting formats, and data governance.

2 RATIONALE - FRAMING THE CHALLENGE

Micro, small, and medium enterprises (MSMEs) represent more than 90 percent of global businesses and are central to inclusive economic growth, job creation, and poverty reduction.[4]

Yet despite their prominence, W-MSMEs, estimated to comprise 38 percent of the global MSME sector, continue to face disproportionate barriers in accessing finance, largely due to the absence of gender-sensitive classifications and robust data systems.[5] To address the global credit gap of USD1.4 to USD1.7 trillion facing W-MSMEs, it is essential to first define and classify them effectively and consistently across countries and data platforms.[6]

2.1 Key concepts and classifications

Defining MSMEs

MSMEs are typically classified based on the number of employees, annual revenue, or asset size. While country definitions vary, global frameworks such as those from the International Finance Corporation and World Bank define micro-enterprises as employing fewer than 10 people and generating annual turnover under USD100,000; small enterprises as having 10-49 employees; and medium enterprises as 50-249 employees with up to USD15 million in revenue.[7] However, these classifications are not standardized across national contexts, leading to difficulties in comparing and aligning data globally.[8]

Defining women-owned and women-led MSMEs (W-MSMEs)

W-MSMEs are enterprises either owned or led by women. A commonly accepted threshold for "women-owned" is 51 percent equity ownership by women, while "women-led" refers to businesses where women serve as CEOs, top managers, or hold key decision-making roles.[9] However, rigid adherence to the ownership threshold excludes enterprises where women are influential leaders but do not meet the ownership cut-off. For instance, Mexico's business registry includes enterprises with significant female board representation and leadership in public procurement initiatives.[10] Excluding such enterprises from formal data systems risks undercounting their contributions and financing needs.[11] The results from the newly developed MSME Data Development Index (DDI) survey conducted in 2025 indicate that gender-specific definitions are far less widespread than general MSME definitions. Among the 47 institutions that responded to this question, 64 percent stated that they do not use a formal definition for W-MSMEs, reflecting misalignment that can result in inconsistent data, poor comparability, and policy blind spots.

Understanding gender-disaggregated data and key challenges

Gender-disaggregated data involves separating data by sex to assess differences in outcomes, access, and participation. In the MSME context, this includes ownership, leadership, access to financial products, loan approvals, digital engagement, and formal registration, all broken down by gender. Such data is critical for policy development and inclusive finance, enabling institutions to design tailored interventions.[12] While gender-disaggregated

[4] World Bank. 2019. MSME Economic Indicators. Available at: https://www.smefinanceforum.org/data-sites/msme-country-indicators

[5] International Trade Centre. 2024. SheTrades Country Profiles. Available at: https://outlook.shetrades.com/

[6] World Bank. 2023. Gender Data Portal. Available at: https://genderdata.worldbank.org/

[7] World Bank. 2019. MSME Economic Indicators, 2019. Available at: https://documents1.worldbank.org/curated/en/873301627470308867/pdf/Micro-Small-and-Medium-Enterprises-Economic-Indicators-MSME-EI-Analysis-Note.pdf

[8] Organisation for Economic Co-operation and Development. 2024. SME Data Lake on SMEs and Entrepreneurship. Available at: https://www.oecd.org/en/data/dashboards/oecd-data-lake-on-smes-and-entrepreneurship.html

[9] SME Finance Forum. 2019. Women SME Finance Data. Available at: https://www.smefinanceforum.org/

[10] Organisation for Economic Co-operation and Development. 2025. Financing SMEs and Entrepreneurs Scoreboard. Available at: https://www.oecd.org/content/dam/oecd/en/publications/reports/2025/04/oecd-financing-smes-and-entrepreneurs-scoreboard-2025-highlights_e7caeca1/64c9063c-en.pdf

[11] International Trade Centre. 2024. SheTrades Country Profiles. Available at: https://outlook.shetrades.com/

[12] Data2X/WFID. 2023. WFID Synthesis Brief. Available at: https://data2x.org/resource-center/wfid-synthesis-brief/

data is often framed around women's inclusion, it is important to collect data for all genders to enable meaningful comparisons and inform equitable policies. Platforms such as the International Monetary Fund (IMF)'s Financial Access Survey and World Bank's Global Findex provide sex-disaggregated indicators on, among other things, account ownership and borrowing behavior, directly informing gender-responsive programming.[13]

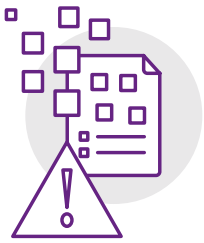
While progress has been made in the collection and usage of gender disaggregated data, various gaps persist as outlined in **Figure 1** below:

FIGURE 1.
KEY CHALLENGES OF GENDER-DISAGGREGATED DATA

Policymakers and practitioners face significant barriers when trying to gather accurate data on women-owned businesses. These challenges create blind spots that hinder effective program design and policy implementation.

LACK OF STANDARDIZED DEFINITIONS

Many countries lack a unified definition of "women-owned or women-led MSME," making cross-country comparisons and data aggregation difficult.

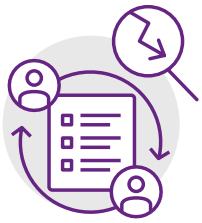


FRAGMENTED DATA ECOSYSTEMS

Data is often siloed across ministries, regulators, financial institutions, and development partners, with limited interoperability or coordination.

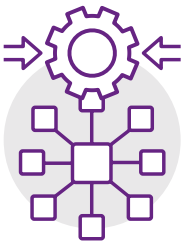
LIMITED SUPPLY-SIDE DATA

Financial Service Providers may not collect sex-disaggregated data at onboarding or during service delivery, especially for informal or micro-enterprises.



LOW INCENTIVES FOR REPORTING

Private sector actors may not see the value in collecting gender data unless mandated by regulators or incentivized through tangible benefits.



CAPACITY CONSTRAINTS

National statistics offices and regulators may lack the technical capacity or tools to collect, analyze, and publish gender-disaggregated MSME data.

PRIVACY & SENSITIVITY CONCERNS

In some contexts, collecting gender data may raise concerns about privacy or cultural sensitivities, leading to underreporting or avoidance.

2.2 Typologies of W-MSMEs

Three key dimensions shape the typologies of W-MSMEs: ownership and control, formality, and size. These classifications influence how enterprises are counted, financed, and included in national economic planning.

By ownership and control

- Women-owned: ≥51 percent owned by women.
- Women-led: Managed or directed by women at the executive or board level.
- Joint Ownership: Shared ownership between men and women.

This classification is critical since policy programs that define W-MSMEs solely based on ownership (e.g. 51 percent rule) may overlook enterprises that are women-led but not majority-owned. For instance, credit and procurement programs in Colombia and Mexico have adopted more inclusive criteria to address this oversight.[14]

[13] World Bank. 2021. Global Findex Database. Available at: https://www.worldbank.org/en/publication/globalfindex

[14] Organisation for Economic Co-operation and Development. 2025. Financing SMEs and Entrepreneurs Scoreboard. Available at: https://www.oecd.org/en/publications/2025/04/oecd-financing-smes-and-entrepreneurs-scoreboard-2025-highlights_e7caeca1.html

By formality

- Formal Enterprises: Officially registered, taxed, and regulated businesses.
- Informal Enterprises: Unregistered, operating outside formal legal and financial systems.

Women-led businesses are often concentrated in the informal sector. For example, Kenya and Bangladesh show how many women-run enterprises are excluded from national datasets due to low registration rates, despite their economic contributions.[15] Informality reduces access to financial services and makes these businesses invisible in policy planning.

This invisibility also poses challenges for supervisors and regulators. When large segments of women-led enterprises operate outside formal registration and reporting systems, it becomes difficult to assess credit exposures and portfolio quality. Using alternative data sources – such as mobile money transactions or digital payment histories – can help improve risk assessment, expand access to finance, and provide regulators with a clearer view of this underserved segment.

By size

- Micro Enterprises: Fewer than 10 employees
- Small Enterprises: 10-49 employees
- Medium Enterprises: 50-249 employees

Most W-MSMEs globally fall within the micro or small enterprise bracket. According to the Asian Development Bank (ADB)'s Asia SME Monitor (2024), women-led businesses tend to be smaller with limited collateral and lower digital integration, which often deter lenders,[16] making gender-disaggregated size classifications vital for developing appropriate credit products and capacity building programs.

[15] World Bank. 2019. MSME Economic Indicators. Available at: https://www.smefinanceforum.org/data-sites/msme-country-indicators

[16] International Trade Centre. 2024. SheTrades Country Profiles. Available at: https://outlook.shetrades.com/

2.3 Significance of classification for gender-responsive data systems

The AFI survey results reveal that gender-responsiveness remains the most underdeveloped aspect in MSME data systems. Out of 54 institutions surveyed, 48 scored 0. Six scored 1, reflecting initial or fragmented efforts, such as collecting isolated gender-specific indicators or expressing interest in tracking gender without institutionalization (see section 4.1.2 for a full breakdown of the DDI survey pilot results).

Even among the small number of institutions that do collect gender data, inconsistent or overly rigid definitions undermine comparability and coverage. Definitions of W-MSMEs vary widely – some require 51 percent ownership by women, others include leadership roles regardless of ownership share. Where definitions are too narrow, many legitimate women-led MSMEs are excluded from counts.

For example, funding or procurement programs that apply a strict 51 percent ownership threshold may miss enterprises where women are CEOs or board leaders but hold a smaller ownership stake. This rigidity leads to systematic undercounting, especially in micro and informal segments.

The absence of consistent, inclusive definitions has two major consequences:

- Data from different countries or institutions cannot be reliably compared, limiting its value for regional or global policy coordination.
- Some women-led enterprises remain ineligible for financing, guarantees, or procurement opportunities that could support their growth.

When robust classification systems are combined with a small set of consistent, high-value indicators, they enable concrete policy and market actions. Under the WE Finance Code,[17] – now implemented in over 30 countries – the five core indicators (number of women-led MSME borrowers, volume of financing, deposits, non-performing loans, and approval rates, all disaggregated by business size) have been used to:

- Structure gender bonds and guarantee schemes, which require historical gender-disaggregated portfolio data.

[17] The Women Entrepreneurs Finance Initiative is abbreviated as We-Fi, while the multi-country data and finance framework is referred to in full as the WE Finance Code.

- Inform regulatory incentives, such as Mexico's lower provisioning requirements for loans to women, based on stronger repayment performance.
- Target public procurement or geographic lending programs, guided by dashboards showing underrepresented segments.

These cases show that gender classification is not just about counting enterprises – it directly shapes financial products, investment flows, and policy incentives. AFI's member-driven frameworks draw on similar principles to help institutions translate gender data into actionable policy and supervisory tools.

Proper classification also enhances consistency, visibility, and effectiveness in the design of targeted solutions. India's Udyam Registration Portal integrates gender-disaggregated fields and supports real-time policy feedback loops, enabling financial institutions to design and disburse credit more equitably. Kenya's FinTech platforms leverage behavioral and mobile usage data to build alternative credit scoring systems for women, showing how nuanced typologies can drive innovation in inclusive finance. These country experiences reinforce the same principle demonstrated by the Women Entrepreneurs Finance Initiative (We-Fi) examples: when classification is clear, inclusive, and tied to relevant indicators, it becomes a powerful enabler of innovation and targeted action.

Classification systems should also extend to informal enterprises, which make up the majority of women-led MSMEs in many low and middle-income countries. Even without formal registration, these businesses often interact with the formal financial system, for example, through microloans, mobile money, or merchant accounts.

By using transaction-level data from regulated financial service providers, mobile network operators, and FinTechs, and linking this to gender via KYC data, national IDs, or machine-learning inference, policymakers can include informal women-led enterprises in aggregate statistics. This improves visibility and policy reach without requiring immediate formalization.

FIGURE 2.
WHY DOES CLASSIFICATION MATTER?

Classification should not be treated as a separate or bureaucratic exercise, but as a core dimension of all MSME indicators – to be disaggregated and analyzed systematically across ownership, size, formality, access, and outcomes. Across countries, definitions of MSMEs and W-MSMEs vary widely, making it difficult to align data and compare outcomes. This diversity underscores the importance of developing harmonized yet flexible classification frameworks. Countries that invest in clear, inclusive, and consistent definitions, paired with relevant indicators, are better positioned to allocate resources effectively and design gender-responsive financing systems.

CLARIFIES DEFINITIONS OF WOMEN-OWNED OR WOMEN-LED ENTERPRISES

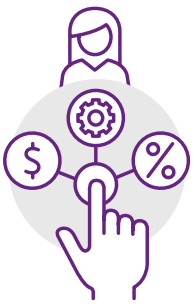
Standardizing definitions ensures consistent data collection on women's micro, small, and medium enterprises (W-MSMEs).
> Misses W-MSMEs if criteria for what counts as "women-owned" or "women-led" are ambiguous.

ENABLES TARGETED PRODUCT & SERVICE DESIGN

Classifying customers by gender and business type supports development of tailored financial products and services.
> Fails to recognize the unique needs of women clients when developing offerings.

IMPROVES POLICY EFFECTIVENESS & MONITORING

Sex-disaggregated data allows more effective monitoring of gender gaps and policy impacts.
> Limits insights into whether financial inclusion policies benefit women.



SUPPORTS RESOURCE ALLOCATION & INCENTIVE DESIGN

Clear classification helps allocate resources like quotas, subsidies, or incentives to advance women's financial inclusion.
> Lacks clarity in directing resources to eligible women or women-led businesses.

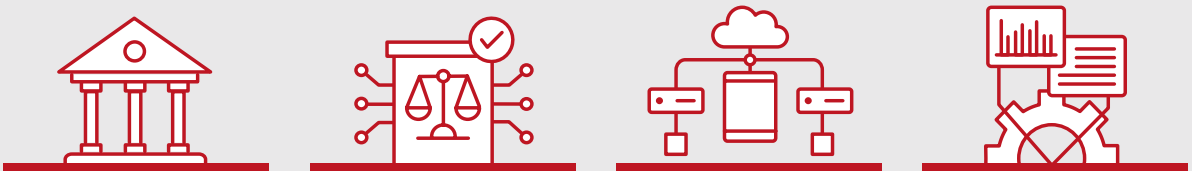
3 BUILDING BLOCKS OF A GENDER-RESPONSIVE MSME DATA ECOSYSTEM

To meaningfully close the gender credit gap and ensure financial inclusion for W-MSMEs, countries must invest in foundational systems that enable the collection, analysis, and application of gender-disaggregated MSME data.

These systems are not created in isolation; as outlined below, they depend on a strong institutional foundation, enabling legal and policy environments, robust digital infrastructure, as well as data use and impact pathways (**Figure 3**).

The experiences of countries such as India, Kenya, Bangladesh, and Mexico demonstrate that when these building blocks are well-coordinated, they can facilitate inclusive and evidence-based policy frameworks for financing W-MSMEs.

FIGURE 3. BUILDING BLOCKS OF A GENDER RESPONSIVE MSME DATA ECOSYSTEM



STEP 1	STEP 2	STEP 3	STEP 4
INSTITUTIONAL FOUNDATIONS	LEGAL & POLICY ENABLERS	INFRASTRUCTURE & PLATFORMS	DATA USE & IMPACT PATHWAYS
Map stakeholder ecosystems	Promote compliance to data protecion and privacy frameworks	Set minimum technical requirements	Build capacity for data interpretation
Define mandates and governance structures	Provide regulatory incentives for gender disaggregated data collection/usage/ sharing	Utilize unified data hubs	Translate data into policy and product design
Establish roles and coordination mechanisms	Embed ethical considerations for gender data governance	Develop interoperability and unique identifiers	Establish feedback and learning loops

STEP 1 - Institutional foundations

Establishing a gender-responsive MSME data ecosystem begins with identifying the relevant institutional actors and establishing clear mandates and coordination mechanisms. AFI members have noted that weak institutional foundations often lead to fragmented stakeholder engagements, undefined mandates for gender data and weak coordination among participating institutions. This, in turn, makes it difficult to track and support women-led enterprises.

Map stakeholder ecosystems

A well-functioning ecosystem involves collaboration between financial sector regulators (e.g. central banks), financial service providers (FSPs), national statistical offices (NSOs), business registries, women's business associations, civil society organizations, and FinTech platforms.

Influence-interest mapping can be used as a strategic tool to analyze and prioritize stakeholders based on how much power or authority a stakeholder has in order to affect decisions, policies, or outcomes (influence) and how much a stakeholder cares about or is affected by the issue, project, or policy in question (interest).

For example, Bangladesh's ecosystem includes contributions from a cross-section of stakeholders who collect and utilize gender-disaggregated credit data. They include the central bank, local microfinance institutions (MFIs), and NGOs. Similarly, India's Ministry of MSMEs operates the Udyam Registration Portal (see case study 1 in chapter 4) with integrated gender indicators, supported through a multi-stakeholder collaboration including commercial banks, credit guarantee agencies, and others.

Mandates and governance structures

Clear institutional mandates, which are essential for defining roles in data collection and ensuring accountability, should be aligned with national gender and financial inclusion strategies. For instance, Kenya's financial regulators are mandated to collect gender-specific indicators for digital credit scoring, as seen in platforms such as M-Shwari and KCB M-Pesa, supported by broader policy mandates for financial inclusion. Governance structures which include inter-agency governance bodies with gender-balanced representation, must define how data is reported, stored, and shared, especially when multiple institutions are involved.

Experience from the Consultative Group to Assist the Poor (CGAP)'s multi-country work suggests that the most effective gender-responsive MSME data systems assign complementary roles across the ecosystem:

Regulator or financial supervisor: holds the legal mandate to request priority MSME indicators (including gender) from regulated institutions; validates and aggregates this data; and, ideally, cross-checks it against national ID and corporate registry databases to ensure accuracy of ownership and leadership information.

1 **FINANCIAL SERVICE PROVIDERS (FSPs)**
Capture gender and MSME size data at the point of onboarding (KYC) and in ongoing transactions; report this data according to the regulator's templates and timelines.

2 **NATIONAL STATISTICS OFFICE (NSO)**
Integrates MSME and gender modules into enterprise surveys, aligns classifications with the regulator's definitions, and supports data comparability over time.

3 **BUSINESS OR ENTERPRISE REGISTRIES**
Record the gender of ownership and leadership at registration, update this information periodically, and make it available, ideally via Application Programming Interface (API), to the regulator and FSP.

4 **CREDIT REGISTRIES OR BUREAUS**
Collect and report credit histories for MSMEs, disaggregated by gender where possible; in some countries, act as sole collectors of credit-side MSME data.

5 **FINTECHS AND OTHER NON-BANK FINANCIAL INSTITUTIONS (e.g. MFIs, cooperatives, MNOs)**
Provide transactional and digital finance data linked to gender and MSME status.

In this "ideal" setup, the regulator acts as the anchor institution, but the mandate to collect and share MSME gender data can also sit with another body (e.g. Ministry of Trade or Economy) as long as strong legal provisions and data-sharing agreements ensure interoperability. The goal is a "plug-and-play" ecosystem in which data is collected once, validated at source, and reused across institutions, thus reducing duplication, improving quality, and lowering reporting burdens on MSMEs.

Roles and coordination mechanisms

Building on these clearly defined mandates, effective gender-responsive MSME data ecosystems depend on structured coordination mechanisms that ensure each actor's role is connected to the others. Without these links, even well-designed mandates risk operating in isolation, creating data silos and missed opportunities for integrated analysis.

Effective gender data ecosystems require a lead institution or inter-agency body to coordinate efforts across sectors. Such collaborations can be guided by memorandums of understanding and data sharing protocol agreements. Coordination ensures that gender-disaggregated data from FSPs, tax agencies, business registries, and civil society organizations is harmonized, avoids duplication, and enables institutional feedback loops with women-owned and women-led enterprises.

Colombia, for instance, coordinates its rural credit programs for women through Finagro, a development finance institution that works closely with ministries, banks, and farmer associations.[18]

RESOURCES AND TOOLS

- Gender Data System Maturity Model
- Sample MoU for Data Sharing Collaboration
- Data Mapping Tool for National Gender Data Ecosystems (We-Fi and Financial Alliance for Women) – provides a structured approach to identify data sources, flows, and responsible institutions within a country's WMSME data ecosystem.

[18] World Bank. 2025. Access to Finance for the Sustainable Transformation of Agrifood Systems. Available at: https://documents1.worldbank.org/curated/en/099042524191511983/pdf/P181242-e2e7e5f0-cb2e-45aa-bb84-e3b861ec76a5.pdf

STEP 2 - Legal and policy enablers

Legal and regulatory frameworks must protect individual privacy while enabling the collection and sharing of gender-relevant MSME data to support inclusive finance. Legal and policy gaps, such as the absence of mandates and incentives for sex-disaggregated reporting, limited enforcement of laws such as data protection laws, low mitigation of ethical risks, and availability of gender sensitive audits are likely to result in data efforts that are optional and often non-prioritized. AFI's Policy Framework on MSME Data Collection provides practical guidance on addressing such gaps, outlining regulatory enablers and reporting templates that have informed the design of this toolkit.

Promote compliance with data protection and privacy frameworks

Gender data systems must comply with national and international data protection and privacy laws, such as the European Union's General Data Protection Regulation (GDPR) and the African Union Convention on Cyber Security and Personal Data Protection (Malabo Convention), to ensure the ethical and lawful management of individual and corporate data.

These frameworks establish standards for data collection, storage, consent, and sharing, emphasizing principles of transparency, proportionality, and data subject rights that are essential for responsible gender data governance. This becomes particularly necessary with the use of financial or behavioral data (such as mobile money transactions) for alternative credit scoring or business diagnostics. Privacy concerns are amplified in contexts where informal women-led enterprises may be more vulnerable to the misuse of data.

Provide regulatory incentives

Governments can accelerate progress by incentivizing FSPs and other actors to disaggregate data by gender. This includes incorporating gender metrics in regulatory reporting requirements, performance-based incentives, or public procurement preferences and tiering reporting requirements for FSPs collecting gender data. A notable example comes from Mexico, where the National Institute of Statistics and Geography (INEGI, the country's autonomous national statistics office) has progressively integrated gender-disaggregated questions into its National Business Survey on Micro, Small, and Medium Enterprises (ENAPROCE) and related data collections. These surveys capture the ownership and management

characteristics of enterprises, including whether they are women-owned or women-led. The resulting datasets allow Mexican authorities to design targeted procurement and credit-support programmes that prioritize W-MSMEs, using evidence rather than assumptions to guide policy.[19]

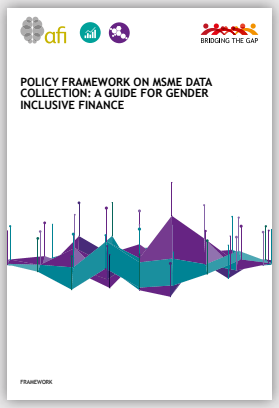
Embed ethical considerations in gender data governance

Ethical data governance ensures that the collection and use of gender-disaggregated data does not reinforce existing inequalities or stigmatize vulnerable groups. This includes obtaining informed consent and data minimization, ensuring that data is not used in a discriminatory way, developing and subscribing to national, regional, and relevant international data governance charters, and engaging women's associations in the governance process. As highlighted by the Data2X WFID (Women's Financial Inclusion Data) initiative, inclusive data governance frameworks help legitimize data systems and build trust among marginalized women entrepreneurs.

Detailed ethical and confidentiality principles for MSME data systems, including consent, proportionality, anonymization, and access traceability, are provided in Annex 4: Ethical and Confidentiality Principles for MSME Data Systems.

RESOURCES AND TOOLS

Gender Data Ethics Checklist



> Policy Framework

[19] Data2X. 2019. Women's Financial Inclusion Data (WFID) Partnership: Mexico Case Study. Available at: https://data2x.org/wp-content/uploads/2019/08/WFID-Mexico-Case-Study_FINAL.pdf

STEP 3 - Infrastructure and platforms

Technology plays a critical role in enabling real-time, gender-tagged MSME data systems that can support financial inclusion efforts at scale. Inadequate digital infrastructure limits real-time data sharing and integration across sources, especially for informal or micro-enterprises run by women, and further perpetuates gender data inequalities in areas with little or no infrastructure.

Unified data hubs

Countries need centralized platforms that consolidate MSME-related data from multiple sources such as business registries, tax authorities, credit bureaus, and financial institutions, creating centralized or federated data repositories that enable real-time access for authorized users. India's Udyam Registration Portal is a successful example, where gender fields are embedded within a unified MSME registry, allowing the state and private sector to target underrepresented groups.

In several countries, national dashboards are now connected directly to financial institutions through secure APIs, enabling automated submission and aggregation of gender-disaggregated MSME indicators. For example, under the WE Finance Code and the Financial Alliance for Women, Bangladesh Bank and the Central Bank of Nigeria have deployed dashboards that pull data directly from FSPs, disaggregated by gender and MSME size, with options to drill down by region or product type. This automation reduces reporting burdens, improves timeliness, and supports near real-time policy monitoring.

To ensure that centralization enhances rather than undermines data reliability, it is essential to incorporate data quality standards and validation protocols within these hubs, preventing the accumulation of inconsistent or unreliable information.

Interoperability and unique identifiers

To track the progress and needs of women entrepreneurs across systems, interoperable platforms promoting cross sector data integration and unique identifiers, such as national ID-linked business registration, are essential. This facilitates integration between business licensing, social protection, and financial services. Bangladesh's dashboard-based reporting system, which connects gender data across rural and urban credit systems, illustrates how interoperability can drive

targeted outreach and bundled services. Interoperability should also be accompanied by a shared data governance framework that defines clear institutional responsibilities, minimizes duplication, and reduces the reporting burden on financial institutions.

Machine learning can be used to infer gender based on the available information where gender fields are missing in legacy datasets. Standard Chartered, for instance, developed a name-recognition algorithm that cross-references client names and ownership data with national registries, achieving 95 percent accuracy across 40 countries, while Morocco's MSME Observatory uses a similar method to tag gender in administrative tax and statistics records. These approaches offer a low-cost, scalable way to enhance gender tagging without requiring full re-onboarding of clients.

Minimum technical requirements

Robust data systems require core technical capabilities including cloud-based storage, secure APIs for data exchange, and analytic tools for real-time reporting and visualization. In addition, standards for data formats, metadata, and APIs should be clearly defined with consideration for accessibility by low resource institutions.

These features are crucial for enabling dynamic dashboards like those in the World Bank's Gender Data Portal and International Trade Centre (ITC)'s SheTrades platform, which provides policymakers with timely insights into gender disparities in MSME financing.

RESOURCES AND TOOLS

Interoperability Assessment Tool

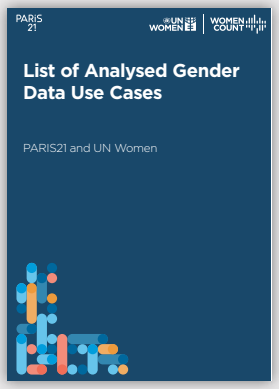
STEP 4 - Activate data use and impact pathways

The value of gender-disaggregated MSME data lies in its use, by policymakers, FSPs, and entrepreneurs, to drive inclusive outcomes. In scenarios where data usage and impact pathways are underdeveloped, data collected remains siloed, unused, or misinterpreted, failing to inform decision-making or close gender gaps in access to finance and support services.

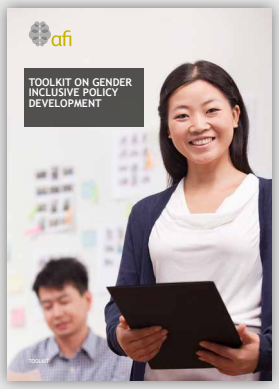
Together, these pillars tackle structural, technical, and strategic shortcomings that inhibit evidence-based, gender-equitable MSME development.

Developing a talent pool of regulators and FSPs who can analyze and interpret data through training on gender data analysis is a key first step. Participatory policy development and product design with women entrepreneurs ensures a higher likelihood of the solutions being fit for purpose. In addition, feedback and learning loops create opportunities for constructive iterations. Over time, institutions can complement these capacity building efforts with periodic assessments to evaluate whether the use of gender-disaggregated information is contributing to narrowing the financing gap for W-MSMEs and strengthening risk management practices.

RESOURCES AND TOOLS



> Gender Data Use Case Library



> Policy Briefing Toolkit

InBrief: Building a National Gender Data Dashboard

Collecting and Using Banking Data on Women Businesses: A How-To Guide for Financial Institutions

4 DATA DEVELOPMENT INDICATOR (DDI). ASSESSING THE MSME DATA ECOSYSTEM

4.1 Using the Data Development Indicator (DDI) to assess institutional maturity on MSMEs data collection

What is the Data Development Indicator (DDI) and what does it measure?

The **Data Development Indicator (DDI)** is a practical tool designed to help institutions assess the maturity of their MSME data systems, with a specific focus on **gender responsiveness**. It evaluates how effectively an institution collects, manages, and uses sex-disaggregated data to inform and guide policies, programs, and financial inclusion strategies.

The DDI is built as a **multi-dimensional maturity model** that does not look at data systems through a single lens but instead breaks down institutional capacity across **four key dimensions:**

1 **DATA COLLECTION**
Are institutions collecting sex-disaggregated data on MSMEs systematically and at scale?

2 **GENDER RESPONSIVENESS**
Are gender considerations integrated into data definitions, indicators, and reporting frameworks?

This dimension remains separate in the DDI for practical purposes, to highlight a widespread gap across institutions. However, gender is not a separate type of data, it is a cross-cutting lens that should be applied across all MSME indicators.

3 **DATA USAGE AND INTEGRATION**
Is the data actually being used for policymaking, programming, or financial service delivery?

4 **GOVERNANCE AND INFRASTRUCTURE**
Are there coordination mechanisms, mandates, and systems in place to ensure data quality, consistency, and sustainability?

This **multi-dimensionality is essential**: it reflects the reality that institutions may be strong in one area (e.g. data collection) but still weak in others (e.g. data usage or coordination). By identifying both strengths and gaps across these dimensions, the DDI supports a **targeted, step-by-step approach to system strengthening** and is designed to accommodate institutions at different levels of data maturity. It does not assume a one-size-fits-all path to maturity, but instead helps tailor improvements based on local context, capacity, and goals.

Each dimension is scored from **0 to 3** and then mapped to one of three institutional maturity levels: **early stage, emerging or advanced**. This enables institutions to identify targeted areas for improvement and to benchmark progress over time.

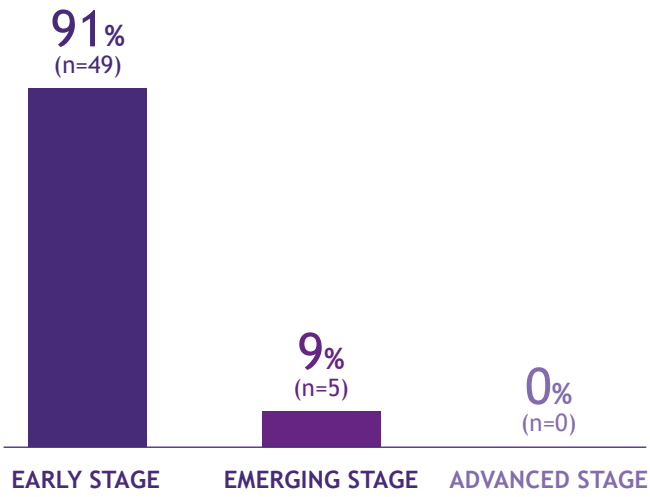
A detailed breakdown of the DDI scoring logic – including the full set of survey questions, indicators, and scoring criteria – is available in **Annex 1**.

How institutions score globally: key trends from the DDI

The DDI was piloted through a survey conducted between March and April 2025, survey conducted to 54 public institutions between across the AFI network – including central banks, financial regulators, national statistics offices, and SME authorities – to assess the maturity of their MSME data systems through a gender lens. The results reveal that the vast majority of institutions are still in the early stages of developing robust MSME data systems.

91% Among the 54 institutions surveyed, the majority fall into the “early stage” maturity category, with 91 percent (49 institutions) scoring between 0 and 3, out of a possible 12. Only five institutions (nine percent) reached the “emerging” category (scoring 4-7), and none qualified as “advanced”.

FIGURE 4.
INSTITUTIONAL MATURITY IN MSME DATA SYSTEMS (DDI PILOT)



This pattern highlights a widespread need for foundational investments in MSME data systems, particularly in integrating **gender considerations**, which remains the weakest dimension.

DIMENSION-SPECIFIC INSIGHTS

1

GENDER RESPONSIVENESS is the least developed area:

89% of institutions scored 0 (no gender-disaggregated MSME data or related indicators).

Only **6 institutions** showed early efforts to integrate gender (e.g. ownership definitions or credit access by sex).

None scored above 1 in this dimension.

2

DETAILED DATA COLLECTION remains limited and narrow:

81% scored 0, showing no structured approach to microdata collection.

10 institutions scored 1, typically those conducting irregular surveys or depending on external data sources.

No institutions scored above 1.

3

DATA USAGE AND INTEGRATION shows more variation:

25 institutions (46 percent) scored 0, indicating no use of MSME data in strategy, policy, or programming.

19 institutions (35 percent) scored 1, typically referencing data occasionally in reports or assessments but without regular or structured application.

10 institutions (19 percent) scored 2, showing moderate integration through dashboards, internal monitoring tools, or program targeting.

4

GOVERNANCE AND INFRASTRUCTURE lags but shows signs of progress:

31 institutions (57 percent) scored 0, indicating no internal coordination structures or data teams related to MSME data.

19 institutions (35 percent) scored 1, typically citing informal or partial structures without a clear mandate.

4 institutions (7 percent) reached a score of 2, reflecting the presence of a dedicated unit, technical staff, or formal internal processes to manage MSME data systems.

None scored 3, which would require institutionalized, system-wide data governance aligned with national strategies.

FIGURE 5. INSTITUTIONAL PERFORMANCE ACROSS DDI DIMENSIONS (DDI PILOT)

These findings illustrate that while **interest in improving MSME data is high**, most institutions are in the **early stages of system development**. Efforts are often siloed, ad hoc, and under-resourced, particularly when it comes to collecting, tracking and supporting **women-led MSMEs**. The next section provides a step-by-step guide for institutions to use the DDI as a self-assessment tool to identify priority areas for action.

How to use the DDI for self-assessment

While the DDI is designed for self-assessment, many institutions may benefit from facilitated sessions – especially during the first application. Capacity building workshops, technical assistance, or peer exchanges can help ensure consistent scoring and promote shared understanding across departments. Chapter 7 of the toolkit outlines options for DDI-related training and technical assistance – including internal workshops, peer support, and facilitated assessments.

Use this four-step guide to apply the DDI in your own institution to identify strengths, gaps, and next steps in building a gender-responsive MSME data system. This can be done as an internal review or a facilitated workshop.

STEP 1 GATHER YOUR TEAM

Bring together staff from departments that collect, manage, or use MSME data (e.g. statistics, gender, IT, policy).

STEP 2 REVIEW YOUR CURRENT SYSTEMS

Use the DDI scoring guide (see Annex 1) to answer the survey questions across four dimensions: data collection, gender responsiveness, data usage and integration, and governance and infrastructure.

STEP 3 ASSIGN SCORES AND IDENTIFY GAPS

Use the scoring logic to assign a score from 0 to 3 for each dimension. This will provide a snapshot of your institutional maturity level (early stage, emerging, or advanced).

STEP 4 PRIORITIZE ACTIONS

Based on your results, identify one or two dimensions to improve. Use the recommendations and case studies in this toolkit (Sections 4.2 and 4.3) for practical entry points.

4.2 Case studies of promising practices

CASE STUDY: GENDER INCLUSIVE FINANCE ROADMAP – NATIONAL BANK OF RWANDA (BNR)

Data collection | Data usage | Governance | Multi-stakeholder collaboration

WHAT THEY DID

The **National Bank of Rwanda (BNR)** launched its **Gender Mainstreaming Strategy (2022-27)**, introducing the **Gender Inclusive Finance Roadmap** to address persistent gender gaps in financial access and use. This initiative emerged in response to **demand-side data from FinScope surveys**, which showed a rise in women's financial inclusion but also revealed significant disparities in women's access to **commercial banking and DFS**.To close this gap, **BNR partnered with the Ministry of Gender and other national stakeholders** to embed **sex-disaggregated indicators** into **national financial surveys** and **routine reporting** mechanisms. The roadmap was complemented by the launch of the **Women's Financial Inclusion Guidelines**, which offer FSPs tools to integrate gender considerations into strategic planning, product design, and service delivery.

WHY IT WORKED

- **Strong institutional leadership:** The initiative was backed by the Governor of the BNR and aligned with national development strategies, ensuring long-term political support.
- **Data-driven motivation:** The roadmap was grounded in findings from the FinScope 2016 survey, which highlighted gaps in formal financial access among women despite rising inclusion rates overall.
- **Practical, targeted tools:** The Women's Financial Inclusion Guidelines provided actionable templates and checklists for FSPs, making gender data operationally relevant.
- **Collaborative implementation:** The roadmap was co-created with support from Access to Finance Rwanda, AFI, and international partners like Data2X and the World Bank, reinforcing capacity and uptake.

WHAT OTHERS CAN LEARN

1. **Address both supply and demand gaps**
   Use demand-side data to diagnose problems and then build supply-side systems to close them.
2. **Align reforms with national priorities**
   Integrating gender into Rwanda's Vision 2050 and Financial Sector Development Plan increased buy-in.
3. **Provide clear tools, not just policies**
   The BNR's guidelines translate gender goals into concrete actions which FSPs can take.
4. **Build multi-level partnerships**
   Cross-ministry and public-private cooperation ensures traction and sustainability.

CASE STUDY: GENDER DATA DIAGNOSTIC – FINANCIAL SECTOR DEEPENING (FSD) KENYA AND CENTRAL BANK OF KENYA (CBK)

Data collection | Data usage | Indicator design | Internal systems | Multi-stakeholder collaboration

WHAT THEY DID

In 2022, FSD Kenya, in collaboration with the Central Bank of Kenya (CBK) and other financial regulators, conducted a comprehensive Gender Data Diagnostic under the Women's Financial Inclusion Data (WFID) initiative. The goal was to assess how gender-disaggregated data was collected, used, and reported across Kenya's financial ecosystem.

The diagnostic mapped supply-side gender data within banks, SACCOs, mobile money providers, MFIs, and Credit Reference Bureaus (CRBs). It confirmed that nearly all FSPs collect sex-disaggregated data for individual customers – but that regulators such as the CBK and the SACCO Societies Regulatory Authority (SASRA) do not collect it routinely.

The motivation stemmed from persistent gender gaps revealed by the FinAccess survey: although formal financial access had reached 83.7 percent of adults in 2021 (up from 26.7 percent in 2006), usage gaps remained. Specifically, only 46.5 percent of women reported using banking products compared to 58.9 percent of men – a 12.4 percentage point gap. Similarly, men were more likely to use digital financial services such as mobile money, despite high overall uptake. These usage gaps highlighted the limits of focusing on access alone.

WHY IT WORKED

- **Data-backed urgency:** The diagnostic used FinAccess and CRB data to show gender gaps in financial service usage despite near-universal access.
- **Ecosystem-wide engagement:** The initiative involved banks, regulators, digital lenders, SACCOs, and CRBs, creating a shared understanding of gaps.
- **Collaborative analysis:** FSD Kenya and CBK conducted joint reviews across datasets, elevating supply-side data as a strategic input.
- **Structured approach:** Applying the WFID methodology provided a recognized, repeatable framework for gender data diagnostics.

WHAT OTHERS CAN LEARN

1. **Identify where sex-disaggregated data exists but isn't used**
   Mapping both routine (e.g. regular regulatory reports) and non-routine (e.g. internal FSP data or ad hoc surveys) collection helps expose blind spots – places where data is available but not leveraged for policy or oversight.
2. **Measure usage, not just access**
   Behavioral insights into how men and women use financial services enable more precise policy and product interventions.
3. **Engage the full data ecosystem**
   Bringing CRBs, mobile operators, and non-bank lenders to the table ensures comprehensive diagnostics.
4. **Use structured tools**
   Standardized diagnostic frameworks like WFID support credibility and replicability.

CASE STUDY: SME OBSERVATORY ENTERPRISE DATA SYSTEM – BANK AL-MAGHRIB AND SME OBSERVATORY

Data usage | Governance | Innovation | Infrastructure | Multi-stakeholder collaboration

WHAT THEY DID

Bank Al-Maghrib (the central bank of Morocco), in collaboration with the **national SME Observatory**, implemented a **centralized enterprise data system** to improve the quality, coverage, and policy relevance of MSME data. The system integrates data from tax authorities, business registries, social security databases, and credit bureaus into a **unified national enterprise database**.

Recognizing that gender is often missing in administrative datasets, the Observatory applied **machine learning techniques** to infer the gender of firm owners based on names and other attributes, allowing for sex-disaggregated analysis of business performance, access to credit, and survival rates.

The integrated system is used not only for **statistical reporting** but also **to inform public programs**. Its data feeds into the design and targeting of MSME policies, including financing schemes, digitalization support, and COVID-19 recovery initiatives.

WHY IT WORKED

- **Institutionalized data-sharing** agreements (MoUs) between multiple agencies ensured regular, automated data flows.
- **Use of machine learning** allowed the team to work around gender data gaps without waiting for perfect datasets.
- **The Observatory's clear policy mandate** created a positive feedback loop between data analysis and program design.
- **Integration of supply- and demand-side data** enhanced accuracy and usability for diverse stakeholders.

WHAT OTHERS CAN LEARN

1. **Institutionalizing data-sharing** across agencies **reduces fragmentation** and improves data quality.
2. **Predictive tools** such as machine learning can **fill critical gaps** when traditional gender data is unavailable.
3. **Connect data systems directly to policymaking:** When observatories actively inform and monitor public programs, their data becomes more relevant to decision-makers.
4. **Leveraging existing administrative datasets** can be **more cost-effective** than creating new surveys from scratch.

CASE STUDY: WE FINANCE CODE COUNTRY APPLICATION – DOMINICAN REPUBLIC

Data collection | Data usage | Indicator design | Internal systems | Multi-stakeholder collaboration

WHAT THEY DID

In 2023, the Superintendency of Banks of the Dominican Republic drew on the WE Finance Code framework to pilot a national initiative on gender-disaggregated MSME data. Working with financial institutions, public agencies, and development partners, they launched a pilot to **systematically collect and share gender-disaggregated credit data for MSMEs**. The experience offers insights for AFI members considering similar approaches within their own regulatory contexts.

The pilot used a **flexible working definition** of women-led MSMEs – including enterprises with **51 percent female ownership, a female CEO or manager**, or at least **30 percent female representation on the board** – to encourage broad participation from financial institutions while aligning with international good practices.

An **API-based data transfer** mechanism was introduced to streamline reporting from participating banks to the Superintendency, improving efficiency and timeliness. The initiative also **linked with the National Statistics Office, Chamber of Commerce**, and **tax authorities** to identify and integrate existing MSME data sources.

Internally, the Superintendency began embedding gender perspectives into its financial inclusion dashboards, expanding the biennial Gender in Banking report to cover MSME indicators. Public transparency was strengthened via the **SIMBAR portal**, which publishes performance reports, including on inclusion metrics.

WHY IT WORKED

- **Strong convening role:** the Superintendency coordinated multiple agencies and FIs despite no formal legal mandate for gender data collection.
- **Practical definitions:** a flexible, inclusive definition allowed harmonization across institutions without delaying the pilot for legal changes.
- **Early use of APIs:** automated data transfer reduced reporting burdens, improved quality, and set the stage for future scalability.
- **Public accountability:** integrating MSME gender data into public dashboards created visibility for both policymakers and the general public.

WHAT OTHERS CAN LEARN

1. **Pilots can unlock momentum**
   Even without a legal mandate, coordinated pilots can establish proof of concept and build stakeholder buy-in.
2. **Definitions matter, but flexibility helps**
   Agreeing on a working definition early enables harmonized collection while leaving space for future refinement.
3. **Technology enables trust and scale**
   API-based reporting improves timeliness, reduces manual errors, and demonstrates commitment to transparency.
4. **Link public and internal systems**
   Publishing data through dashboards and portals helps institutionalize gender indicators and maintain public accountability.

CASE STUDY: UDYAM REGISTRATION PORTAL – INDIA

Data collection | Data usage | Indicator design | Infrastructure | Multi-stakeholder collaboration

WHAT THEY DID

Launched in 2020 by the Ministry of Micro, Small, and Medium Enterprises (MSME), the **Udyam Registration Portal** provides a **centralized digital platform for the formal registration of MSMEs in India**, streamlining enterprise access to government schemes, finance, and support services. As a digital-first platform, Udyam serves as a **unified national database** for policy design, enterprise verification, and impact monitoring.

The platform is notable for **integrating sex-disaggregated data fields during registration**, enabling **real-time tracking of W-MSMEs**. Dashboards use gender-specific filters to display geographic and sectoral breakdowns of W-MSME performance. This data underpins gender-responsive programs such as credit guarantees, capacity building, and entrepreneurship support.

Udyam was built through a broad **collaborative process** that included the National Statistical Office, banks, industry associations, state governments, and women's business chambers. The result is a **dynamic data ecosystem** that supports **loan validation, policy coherence, and targeted financing**.

WHY IT WORKED

- **Early integration of gender fields** into enterprise registration created a **powerful gender-disaggregated dataset** at minimal cost.
- **Real-time data dashboards** allowed for **performance monitoring** of W-MSMEs across states and sectors.
- **Cross-institutional collaboration** ensured system relevance, technical alignment, and widespread uptake.
- **Linkages** with subsidy programs and loan verification processes made **the data operationally useful** for FSPs and the government alike.

WHAT OTHERS CAN LEARN

1. **Unified data platforms** improve policy efficiency and enable targeted support for underrepresented groups.
2. **Involving a wide set of actors early** – statistics offices, banks, local governments – improves data uptake and sustainability.
3. **Embedding gender data at the point of registration** is a low-cost, high-impact practice.
4. Platforms like Udyam can **serve both regulatory and service delivery goals** through smart integration.

4.3 Emerging patterns and lessons

The following lessons synthesize peer learning across AFI members and partners, highlighting pathways for system strengthening.

1 Policy mandates and alignment drive sustained progress

Across the case studies, strong institutional mandates, whether through national strategies (Rwanda), legal mandates (India), or central bank leadership (Morocco), proved essential to mainstreaming gender-disaggregated data.

2 Gender responsiveness requires both demand- and supply-side data

Several institutions, including the National Bank of Rwanda and the Central Bank of Kenya, demonstrated the value of integrating demand-side diagnostics (e.g. FinScope, FinAccess) with supply-side administrative data. This dual approach not only reveals gender gaps more accurately but also guides more effective interventions.

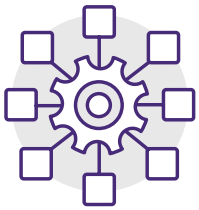
3 Early, low-cost wins can unlock momentum

India's inclusion of gender fields in MSME registration and the Dominican Republic's pilot API-based reporting are examples of relatively simple but high-impact practices. These "low-barrier" interventions can create visibility and buy-in while laying the foundation for more ambitious system reforms.

4 Data use must be designed from the start

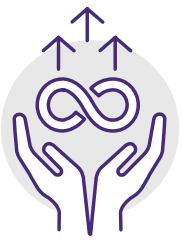
Where data collection was directly linked to policymaking, such as Morocco's integration of data into COVID-19 recovery and SME support programs, data systems had higher relevance and staying power. Tools such as dashboards (India) or public reports (Dominican Republic) not only promote transparency but also help convert data into actionable insights.

5 Interoperability and coordination are essential but hard-won



Effective MSME data systems depend on collaboration across regulators, ministries, banks, and statistical bodies. Case studies from Morocco, the Dominican Republic, and Rwanda underscore the importance of establishing data-sharing agreements, standardizing templates, and building a shared language for gender definitions and MSME classifications. However, these processes often face institutional silos, legal barriers, and misaligned data standards, making sustained coordination a long-term investment rather than a quick win.

6 Definitions matter and flexibility helps



Standardized definitions of "women-led MSMEs" remain a barrier in many countries. The Dominican Republic case illustrates how working definitions (e.g. based on ownership or leadership thresholds) can be pragmatically applied even in the absence of legal codification, allowing systems to move forward while enabling comparability.

CASE STUDY: STANDARD CHARTERED – AI GENDER TAGGING FOR LEGACY MSME PORTFOLIOS

Data collection | Data usage | Indicator design | Infrastructure | Innovation

WHAT THEY DID

To address gaps in historical gender data, Standard Chartered developed an **AI-powered name-recognition algorithm** to retroactively tag the gender of MSME clients in its legacy portfolios.

The process combines:

- Internal client databases, including names and identifiers (e.g. titles such as Mr. and Ms.).
- Cross-referencing with national firm registry data, where accessible.
- Country-specific name lists and linguistic rules to predict gender.

The system outputs a male, female, and unknown tag for each MSME client. Across **40 countries**, the model achieved **95 percent accuracy** in correctly identifying gender, enabling the bank to apply gender-disaggregated analytics to its full lending portfolio without having to re-contact every client, while meeting increasing demands from regulators and investors for gender-disaggregated reporting.

WHY IT WORKED

- **Low-cost solution for legacy gaps:** No need for resource-intensive manual data collection or client re-onboarding.
- **Scalable across markets:** Adaptable to different languages and national contexts by using country-specific name lists.
- **Improved data usability:** Enabled the bank to generate time-series gender-disaggregated indicators for product design, risk analysis, and reporting.
- **Compatibility with external systems:** Gender tags can be integrated into dashboards and reporting templates for regulators or industry frameworks.

WHAT OTHERS CAN LEARN

1. **AI can be a bridge technology**
   Machine learning can fill gender data gaps quickly, making legacy portfolios “analysis-ready” while institutions improve front-end data capture.
2. **Accuracy requires local adaptation**
   Country-specific name dictionaries and registry cross-checks significantly improve results.
3. **Backfilling builds the business case**
   Once gender data is available, institutions can uncover underserved segments and justify targeted products or investment instruments.
4. **Transparency matters**
   Documenting the methodology and accuracy rates helps build trust in AI-generated gender tags among regulators and partners.

CASE STUDY: WE FINANCE CODE – MULTI-COUNTRY GENDER DATA AND FINANCE

Data collection | Data usage | Governance | Indicator design | Internal systems | Multi-stakeholder collaboration

WHAT THEY DID

Introduced in 2023, the WE Finance Code illustrates a multi-partner effort to close finance and data gaps for women-led MSMEs. Although externally led, its indicator framework is broadly aligned with AFI's work on gender-inclusive finance and offers useful lessons for AFI members developing their own national approaches. Endorsed by all major multilateral development banks – the IFC (International Financial Corporation), World Bank, AfDB, ADB, IsDB, IDB, and EBRD – the Code is implemented country-by-country through a national coalition coordinated by the central bank, a bankers' association, or another designated aggregator.

Participating financial institutions commit to:

- Appoint a senior-level champion for women's markets.
- Collect and report gender-disaggregated MSME data.
- Take concrete actions to close financing gaps (e.g. launch new products, gender bonds, or targeted training).

On the data side, all signatories report five mandatory indicators, disaggregated by gender and MSME size: the number of women-led MSME borrowers; volume of financing to them; level of deposits; non-performing loans (NPLs); and loan approval rates.

Optional indicators, such as women in leadership, pay gaps, and climate finance, can be added. Aggregators are supported with API-linked dashboards and analytics tools. By 2025, the Code was active in 31 countries, with over 300 FIs participating, while 12 countries had already launched formal data collection under the framework.

WHY IT WORKED

- **Global standardization with local flexibility:** The five indicators were developed through a year-long consultation with MDBs, regulators, and FIs, and align with Organisation for Economic Co-operation and Development (OECD) and IMF formats, enabling cross-country comparability while allowing local adaptations.
- **Multi-stakeholder buy-in:** National coalitions typically secure participation from 70-80 percent of the financial system, creating broad coverage and peer pressure to maintain quality.
- **Integrated technical support:** Aggregators receive a toolkit, dashboard templates, and API specifications to automate reporting and minimize manual data handling.
- **Policy and market relevance:** Data has been used to structure gender bonds and guarantees (requiring historical sex-disaggregated data), inform incentive schemes (e.g. Mexico's reduced provisioning for women's loans), and target underserved regions or sectors.

WHAT OTHERS CAN LEARN

1. **Start with a small, high-value indicator set**
   A limited number of mandatory indicators makes participation feasible while ensuring policy and market impact.
2. **Combine commitments with tools**
   Reporting requirements work best when paired with dashboards, templates, and technical assistance to ensure data is usable.
3. **Anchor in a credible aggregator**
   Central banks or bankers' associations can act as trusted conveners to align definitions and reporting formats.
4. **Use data to drive action**
   Linking data to product design, investment instruments, and regulatory incentives creates a tangible return on data collection

7 Technology can bridge legacy data gaps

The Standard Chartered AI gender tagging case shows that technology can help institutions quickly backfill missing gender data in legacy portfolios, enabling them to meet reporting requirements and unlock analytics without re-contacting clients. Machine learning and name-recognition techniques can be a cost-effective bridge while front-end systems are upgraded to consistently capture gender fields.

8 Institutions are moving from collection to strategy

These cases show a shift from passive data collection to intentional data strategies. Whether through diagnostics (Kenya), roadmaps (Rwanda), or full integration into enterprise systems (Morocco), institutions are increasingly positioning gender-disaggregated MSME data as a strategic enabler of inclusive financial and economic development.

These eight lessons offer transferable insights drawn from diverse institutional experiences. **But where should institutions begin? The answer depends on their current level of data system maturity.**

The table below links the key case study practices to the **three overall DDI maturity levels**, while also showing how these levels manifest across **the four DDI dimensions: D1 – Data collection, D2 – Gender responsiveness, D3 – Data usage and integration, and D4 – Governance and infrastructure**. This approach reflects the reality that an institution may be "early stage" in one dimension and "emerging" or "advanced" in another and highlights that improvement pathways can be targeted by dimension rather than solely by overall maturity level.

TABLE 1. WHERE TO START? BEST PRACTICES BY DDI LEVEL

DDI maturity level	What this means (by dimension)	Suggested entry points	Relevant case studies
Early stage	**D1:** No systematic collection of MSME gender data. **D2:** Gender not integrated into indicators and definitions. **D3:** Limited or no use of MSME data in strategy or policy. **D4:** No coordination platform or mandate.	• Add gender fields to MSME registration or surveys • Develop a working definition for W-MSMEs • Run a gender data diagnostic • Pilot small-scale data reporting with FSPs	• India (gender at registration) • WE Finance Code – Dominican Republic pilot (flexible definitions, API-based reporting) • Kenya (diagnostic approach)
Emerging	**D1:** Some sex-disaggregated MSME data collected, but partial coverage. **D2:** Gender definitions used inconsistently; some targeted indicators. **D3:** Data occasionally used for program design or dashboards. **D4:** Partial coordination via MOUs; basic governance structures.	• Create dashboards for internal or public use • Formalize inter-agency data-sharing agreements • Use data to design or tune support programs • Align indicators with national priorities	• Rwanda (survey + reporting mandate) • Morocco (multi-source integration) • WE Finance Code – multi-country framework (five core indicators)
Advanced	**D1:** Comprehensive MSME data (formal + informal) collected and integrated across sources. **D2:** Broad gender indicator set embedded in systems; routinely updated. **D3:** Real-time dashboards and analytics inform policies and products. **D4:** Institutionalized governance; well-resourced units manage MSME data.	• Expand gender data into monitoring and strategy • Use real-time dashboards for transparency and planning • Strengthen cross-sectoral governance bodies • Mentor peer institutions in the region	• Rwanda (national roadmap) • Morocco (data-policy integration) • Standard Chartered (AI-enabled gender tagging for legacy portfolios)

5 INDICATOR FRAMEWORK AND DATA ARCHITECTURE

5.1 Tiered indicator framework

Purpose

This section provides a structured set of indicators to guide institutions in building or strengthening their MSME data systems. The framework includes both **core indicators** and **enriched indicators**, organized across key domains such as ownership, finance, business characteristics, and digital inclusion. In addition to supporting the measurement of financial inclusion, the indicator framework can also help regulators identify credit concentrations or differentiated portfolio behavior among W-MSMEs – information that can be useful for macroprudential analysis and supervisory planning.

Tiered approach

The indicator framework is divided into two tiers:

Tier 1 - core indicators : These are the basic set of indicators required to build visibility into the MSME landscape, including women- and men-led enterprises. They are designed to be:

- **Feasible** for most institutions to collect (via administrative or survey data)
- **Aligned** with international practices
- **Actionable** for entry-level analysis and inclusion policies

Tier 1 also incorporates the five core indicators of the WE Finance Code covering the number of borrowers, volume of financing, level of deposits, non-performing loans (NPLs), and loan approval rates. Including these ensures alignment with OECD, IMF, and multi-country reporting formats, and enables institutions to contribute to both domestic and global monitoring efforts.

Tier 1 (core) indicators are highly recommended for institutions at **any DDI level**, particularly those in the **early stage** or e**merging phases**.

Tier 2 (enriched) indicators offer deeper insights into MSME inclusion, exclusion, and outcomes. They are useful for institutions that:

- Already collect or integrate multiple data sources
- Are at an **advanced DDI level**
- Are engaging in more **targeted policymaking** (e.g. public procurement for W-MSMEs or FinTech oversight)

While Tier 2 indicators are not essential for all, they are critical for **expanding policy relevance** and **embedding gender responsiveness** across the data system.

List of key indicators

The table below presents the recommended Tier 1 (core) and Tier 2 (enriched) indicators for MSME data systems. This toolkit does not treat gender as a separate data type, but as a cross-cutting dimension essential for inclusive analysis and policymaking. Tier 1 provides the foundational visibility needed for policy action, while Tier 2 expands into deeper, more specialized areas that require stronger infrastructure and coordination. Full definitions for each indicator are provided in Annex 1.

Domain	Indicator	Tier
Ownership and leadership	% ownership by gender	Tier 1
	Gender of manager/CEO	Tier 1
	Age of owner	Tier 2
	Education level of owner	Tier 2
Business profile	MSME size classification (micro, small, medium)	Tier 1
	Business registration/formality status	Tier 1
	Business sector (e.g. trade, agriculture)	Tier 1
	Geographic location (region, rural/urban)	Tier 1
	Age of business/Registration date	Tier 2
Access to finance - including WE Finance Code core indicators	Number of MSME borrowers (WE Finance Code core indicator 1)	Tier 1
	Volume of financing to MSMEs (WE Finance Code core indicator 2)	Tier 1
	Level of deposits for MSMEs (WE Finance Code core indicator 3)	Tier 1
	Non-performing loans (NPLs) for MSMEs (WE Finance Code core indicator 4)	Tier 1
	Loan approval rates for MSME applications (WE Finance Code core indicator 5)	Tier 1
	Loan application status	Tier 1
	Loan terms (interest rate, maturity, collateral)	Tier 1
	Use of savings or deposit products	Tier 2
	Use of alternative finance (e.g. microfinance, FinTech)	Tier 2
	Credit bureau presence/credit history	Tier 2
	Number of financial accounts or services used	Tier 2
Digital access	Use of DFS (e.g. mobile money, online platforms)	Tier 1
	Access to the internet/smartphone for business purposes	Tier 2
Public procurement	Participation in public contracts or tenders	Tier 2
	Value or % of contracts awarded	Tier 2
Business outcomes	Revenue or sales trends	Tier 2
	Business survival/exit rate	Tier 2
	Employment created	Tier 2

5.2 Prioritization matrix

Prioritizing indicators based on feasibility and relevance

Not all institutions will be able to collect all indicators at once. The goal of this toolkit is to help institutions **prioritize** the indicators that are **most feasible, relevant, and actionable**, given their current capacity. This prioritization is guided by three factors:

- **Feasibility:** How easy the indicator is to collect (administrative vs. survey-based, complexity, cost)
- **Relevance:** How useful the indicator is for decision-making and inclusion-oriented policies
- **DDI maturity level in Dimension 1 (Data Collection):** An institution's ability to collect, manage, and process MSME data

AFI members can prioritize indicators according to their data maturity and policy relevance.

- Institutions at an **early stage (score 0-1)** should focus on **Tier 1 indicators** that are simple, high-value, and achievable.
- Institutions at an **emerging level (score 2)** may begin **adding Tier 2 indicators** that fill gaps in disaggregation or expand analytic depth.
- Institutions at an **advanced level (score 3)** can **expand into a broader range of Tier 2** indicators and explore complex linkages or proxy data sources.

The quadrant matrix below provides guidance on which indicators to prioritize, based on where your institution stands and where it wants to go.

A 2x2 quadrant matrix for prioritization

A 2x2 grid will be used where:

- **Feasibility** = based on how easy it is to collect the data (linked to DDI Dimension 1: Data Collection)
- **Relevance** = based on policy value and stakeholder interest

	Low relevance *(Useful but not central)*	**High relevance** (Critical for policy design or monitoring)
High feasibility (e.g. registry, admin data)	**NICE TO HAVE** Can be added over time or for specific programs	**QUICK WINS** Core indicators that should be prioritized even at early stages
Low feasibility (e.g. surveys, linked data)	**DEPRIORITIZE** Not urgent or may not justify investment at current stage	**STRATEGIC PRIORITIES** High-value indicators to plan for as capacity increases

HOW TO USE THIS MATRIX?

- Institutions at an **early stage (DDI 1 score 0-1)** should focus on the **top-right quadrant (quick wins)**.
- **Emerging institutions (DDI 1 score 2)** may explore both **top-right** and **bottom-right (strategic priorities)**.
- **Advanced institutions (score 3)** can address **all quadrants** except the bottom-right (deprioritize).

QUICK WINS
(top-right: high feasibility, high relevance)

These indicators are both essential and achievable. They are typically available through existing administrative data or simple survey modules, and they provide immediate value for MSME policy and gender analysis. Institutions at any DDI level should start with these indicators. They form the foundation of Tier 1 and should be the first set of indicators to prioritize.

STRATEGIC PRIORITIES
(bottom-right: low feasibility, high relevance)

These indicators are highly valuable but may require additional capacity or data integration. They are ideal for institutions that:

- Are at an emerging or advanced DDI level.
- Have medium- to long-term plans for system strengthening.

These indicators should be part of an institution's medium-term roadmap, especially when aligned with policy priorities (e.g. procurement, credit quality, or business performance).

NICE TO HAVE
(top-left: high feasibility, low relevance)

These indicators are easy to collect but may not currently be central to MSME policy discussions. They can be useful for:

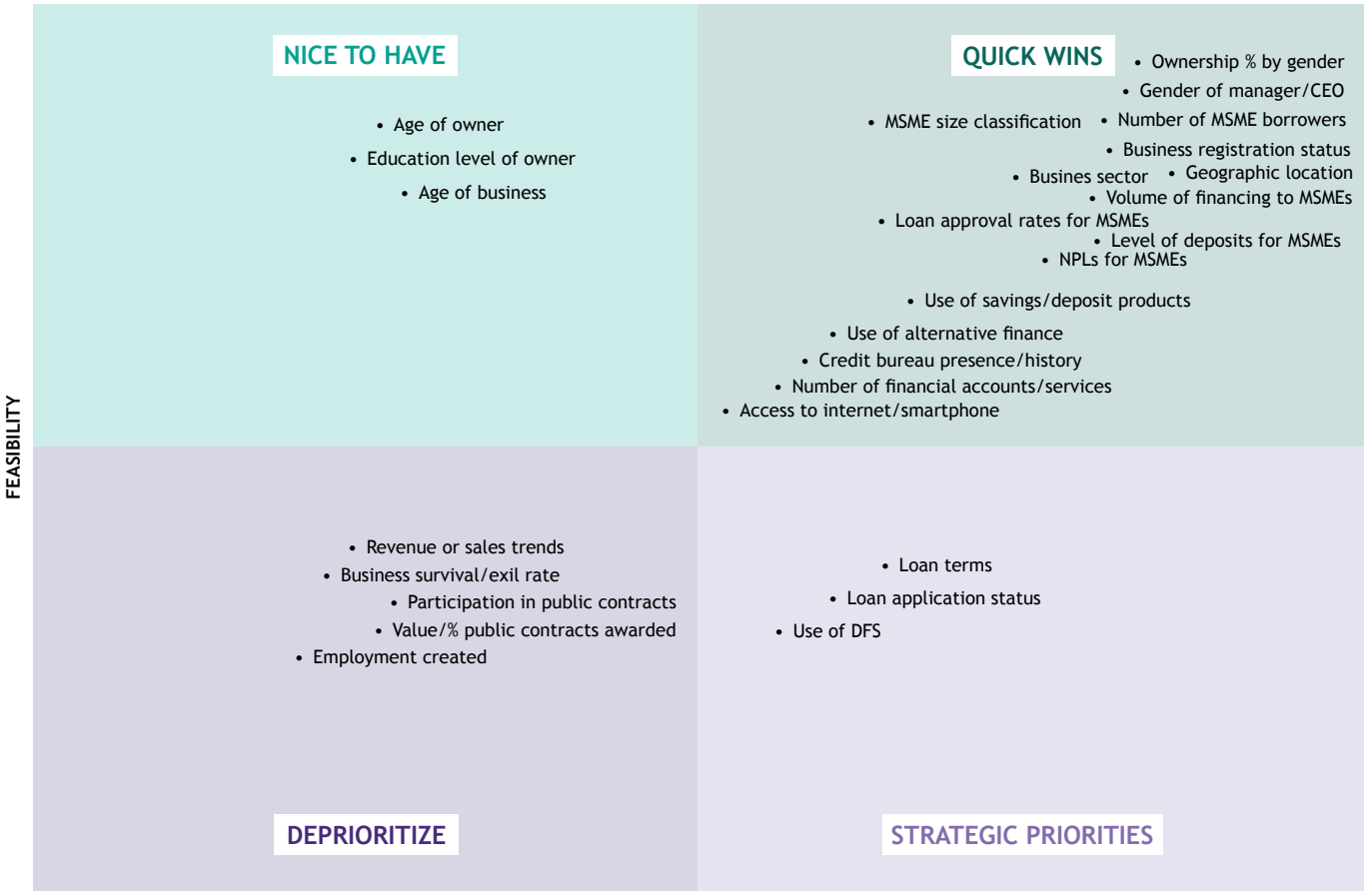
- Sectoral monitoring
- Research
- Supplementing core indicators

Institutions may include them if they are already available or if they address a specific policy niche, but they are not essential to start.

DEPRIORITIZE OR REVISIT LATER (bottom-left: low feasibility, low relevance)

These indicators are difficult to collect and low in immediate value. They are not recommended for standard monitoring systems, even for advanced institutions. Use only for specific projects, or revisit as capacity and priorities evolve.

DDI PRIORITIZATION MATRIX - TIER 1 & TIER 2 INDICATORS



Mapping Tier 1 indicators to DDI dimensions

The table below maps each Tier 1 indicator to the DDI dimensions where it is most relevant. This helps institutions identify indicators that contribute not only to **D1 - Data collection**, but also to **D2 - Gender responsiveness**, **D3 - Data usage and integration**, and **D4 - Governance and infrastructure**.

Domain	Indicator	D1	D2	D3	D4	Notes
Ownership and leadership	% ownership by gender	✓	✓			Core gender-disaggregated ownership metric
	Gender of manager/CEO	✓	✓			Captures leadership beyond ownership by identifying the gender of the enterprise's principal decision-maker (e.g. CEO or manager).
Business profile	MSME size classification	✓		✓	✓	Key for segmentation and policy targeting
	Business registration/ formality status	✓		✓	✓	Supports integration with registries/licensing
	Business sector	✓		✓		Enables sectoral policy/ program analysis
	Geographic location (region, rural/urban)	✓		✓		Supports geographic targeting
Access to finance (incl. WE Finance Code core indicators)	Number of MSME borrowers (WE Finance Code 1)	✓	✓			
	Volume of financing to MSMEs (WE Finance Code 2)	✓	✓	✓		
	Level of deposits for MSMEs (WE Finance Code 3)	✓	✓	✓		
	Non-performing loans (NPLs) for MSMEs (WE Finance Code 4)	✓	✓	✓		Risk and product design relevance
	Loan approval rates for MSMEs (WE Finance Code 5)	✓	✓	✓		
	Loan terms (interest rate, maturity, collateral)	✓		✓		Product design and regulatory review
	Loan application status	✓		✓		
Digital access	Use of DFS (e.g. mobile money, online platforms)	✓		✓		

This mapping allows institutions to prioritize indicators not only by their importance for data collection (D1), but also by their strategic value across other DDI dimensions. Indicators that touch multiple dimensions, such as the WE Finance Code core indicators or MSME size classification, often deliver higher returns on investment because they can simultaneously improve gender responsiveness (D2), support evidence-based policymaking (D3), and strengthen governance and interoperability (D4). Institutions can use this table alongside the prioritization matrix in 5.2.2 to select indicators that are both foundational and high impact.

5.3 Data source mapping

Purpose

Once institutions identify which indicators they want to prioritize, the next step is to understand **how to obtain the data**, and whether it needs to be collected at all.

Not all data must be collected from scratch. In many cases, institutions can make strategic use of **existing administrative data, public records, or third-party datasets** to efficiently fill gaps.

This section provides a mapping of:

- **Potential data sources** for each indicator (e.g. tax records, credit bureaus, MSME registries)
- **Collection methods**, such as administrative capture, surveys, FinTech-generated data, or telecom sources (MNOs)
- And the **key disaggregation dimensions** to consider (e.g. gender, business size, region).

By linking each indicator to realistic and available data sources, this mapping supports institutions in:

- Designing cost-effective data strategies
- Identifying opportunities for **data sharing and integration**
- Avoiding duplication and over-surveying

When identifying data sources, institutions should also assess their reliability and temporal consistency, especially when information comes from multiple entities such as tax authorities, credit bureaus, or FinTech platforms. Establishing reconciliation or cross-validation mechanisms between administrative, financial, and statistical records helps reduce duplication and discrepancies, improving comparability and overall data quality. From a supervisory perspective, integrating these data sources can also reveal credit exposure concentrations and correlations between risks, supporting both financial stability analysis and inclusion objectives.

Lessons from recent initiatives show that starting with existing data sources and then layering low-cost, targeted collection methods often yields faster results than launching large new surveys. Where gender fields are missing, technology-assisted methods, such as AI-based name-recognition algorithms used by Standard Chartered and Morocco's MSME Observatory, can help fill gaps without overhauling core systems.

Avoiding duplication and survey fatigue

While data gaps are real, so is the risk of **over-collecting data that already exists**. Repeating data collection efforts can:

- Waste institutional resources
- Create inconsistencies across datasets
- Burden MSMEs (especially women-led firms) with **repetitive surveys and unclear outcomes**

Institutions are encouraged to first **review existing administrative and partner data sources** before launching new collection efforts as **better coordination and smarter data reuse** lead to more efficient, sustainable, and gender-responsive data ecosystems.

This principle is especially relevant for institutions at **early and emerging DDI levels**, as they can often gain major insights by simply connecting or repurposing existing data streams.

Low-cost and technology-assisted methods for filling data gaps

Many institutions face budgetary and capacity constraints that make large-scale, dedicated MSME surveys difficult to implement. In such contexts, low-cost and technology-assisted approaches can help close data gaps quickly while laying the foundation for more systematic collection in the future.

Optimizing existing data sources

Before investing in new collection efforts, institutions can maximize the value of existing administrative datasets (such as business registries, credit bureau records, or social security data) by improving interoperability and linking datasets. Where gender

fields are missing, AI-based name-recognition algorithms can infer the likely gender of business owners or managers using first names. **Standard Chartered** applied this method to its historical lending data to enhance gender visibility, while Morocco's MSME Observatory piloted similar tools on registry data to produce gender-disaggregated counts for policy planning.

Capturing informal MSMEs

Informal enterprises, many of them women-led, are often absent from official datasets. KIIs from **CGAP** and **WE Finance Code** highlight cost-effective methods to bring them into view, including:

- Adding short informal sector modules to ongoing national household or labor force surveys.
- Community-based mapping and local enumerator approaches.
- Leveraging mobile network data or geospatial analysis to detect business activity.

Integrating with ongoing surveys and activities

Instead of launching standalone MSME surveys, institutions can "piggyback" on existing data collection efforts. This includes inserting targeted MSME and gender questions into ongoing national surveys, collaborating with chambers of commerce or sector associations to collect basic data during membership renewals, or using digital enumerator tools to reduce field costs.

How to source your indicators: linking to existing data and methods

The table below links each Tier 1 indicator to its most likely **secondary data source**, the recommended **collection method** where needed, and the **disaggregation points** to consider when reporting. This helps institutions understand where data is most likely to come from and how it should be structured for meaningful use.

Domain	Indicator	Potential secondary sources	Data collection methods	Disaggregation dimensions
Ownership and leadership	% of ownership by gender	Business registries, surveys, tax records, licensing data	Administrative data, surveys, AI-based name-recognition algorithms	Gender, MSME size, region
	Gender of manager or CEO	Business registries, surveys, association membership lists	Administrative data, surveys, AI-based name-recognition algorithms	Gender, MSME size, region
Business profile	MSME size classification	Business registries, tax filings, surveys	Administrative data, surveys	MSME size, region
	Business registration or formality status	Business registries, licensing authorities	Administrative data, surveys, community-based mapping, mobile data, and informal sector modules in household surveys	Gender, MSME size, region
	Business sector	Business registries, sector associations, surveys	Administrative data, surveys, community-based mapping, mobile data, and informal sector modules in household surveys	Gender, MSME size, region

Domain	Indicator	Potential secondary sources	Data collection methods	Disaggregation dimensions
	Geographic location (region, rural or urban)	Business registries, surveys, geospatial data	Administrative data, surveys, community-based mapping, mobile data, and informal sector modules in household surveys	Gender, MSME size, urban/rural
Access to finance	Number of MSME borrowers (WE Finance Code core indicator 1)	Credit bureaus, financial institutions, central bank reporting	Administrative data, financial sector reporting	Gender, MSME size
	Volume of financing to MSMEs (WE Finance Code core indicator 2)	Credit bureaus, financial institutions, central bank reporting	Administrative data, financial sector reporting	Gender, MSME size
	Level of deposits for MSMEs (WE Finance Code core indicator 3)	Financial institutions, central bank reporting	Administrative data, financial sector reporting	Gender, MSME size
	NPLs for MSMEs (WE Finance Code core indicator 4)	Credit bureaus, financial institutions, central bank reporting	Administrative data, financial sector reporting	Gender, MSME size
	Loan approval rates for MSMEs (WE Finance Code core indicator 5)	Credit bureaus, financial institutions, central bank reporting	Administrative data, financial sector reporting	Gender, MSME size
	Loan terms (interest rate, maturity, collateral)	Financial institutions, central bank reporting	Administrative data, financial sector reporting	Gender, MSME size
	Loan application status	Financial institutions, central bank reporting	Administrative data, financial sector reporting	Gender, MSME size
Digital access	Use of DFS (e.g. mobile money, online platforms)	Payment service providers, mobile network operators	Administrative data, surveys	Gender, MSME size, region

6 INTERPRETATION: MAKING DATA USEFUL

6.1 Purpose and scope

This chapter provides **basic guidelines for visualizing DDI results and MSME data**, drawing on proven practices from partner institutions and initiatives. It is not intended as a comprehensive manual on statistical analysis or data visualization, but rather offers practical considerations for presenting gender-responsive MSME data in ways that are clear, actionable, and adapted to different audiences.

Effective data visualization is essential for:

- Supporting better policy and program decisions through timely, accessible insights.
- Enabling stronger advocacy by making gender disparities visible and understandable.
- Communicating findings clearly to diverse stakeholders – from technical analysts to policymakers, financial institutions, and the public.

The guidance in this chapter builds on real-world examples from We-Fi, central banks, SME observatories, and other partners who have developed dashboards, reports, and interactive platforms to track and share MSME and gender data. These examples illustrate both the potential and practical steps involved in transforming raw data into tools for decision-making.

6.2 Visualizing DDI results

The Data Development Index (DDI) is most useful when presented in a way that highlights both **overall progress** and **specific strengths and gaps** across its four dimensions (D1–D4). Visualizations should make it easy for different audiences, from policymakers to technical analysts, to quickly interpret results and identify priority actions.

As the DDI framework is tested and refined, it will be important to maintain a consistent structure for visualizations across countries, and over time, to enable valid comparisons and tracking of progress. When disseminating results, institutions should also include brief methodological notes explaining the meaning of each level, the data sources used, and the scoring criteria, to ensure transparency and accessibility for non-technical audiences.

Overall vs. dimension-level scores
While an overall DDI score provides a simple benchmark, it can mask variations across dimensions. For example, an institution may have a high score for "Data collection" (D1) but a low score for "Governance and infrastructure" (D4). Presenting results by dimension helps target interventions where they are most needed.

Example: In **WE Finance Code's** mock-ups, composite scores are displayed both as an overall index and by dimension in spider or radar charts, with color-coding applied to distinguish performance levels. In the context of the DDI, this same approach can be adapted to use maturity-level color coding (Foundational, Emerging, Established) to guide interpretation.

Recommended formats

- Spider or radar charts to compare dimension-level scores for a single country or institution.
- Heatmaps to show scores across multiple institutions or countries in a matrix format.
- Maturity-level color coding applied to bars, radar chart areas, or heatmap cells to indicate DDI levels (Foundational, Emerging, Established).

6.3 Principles for analyzing and aggregating MSME data

Disaggregate first
Always analyze data disaggregated by gender, MSME size, location, and other relevant dimensions before aggregating results. Aggregation without disaggregation can obscure disparities, particularly those affecting women-led or informal enterprises.

Example: In Morocco's SME Observatory, gender and size breakdowns revealed that women-led micro-enterprises had significantly lower credit approval rates than men-led micro-enterprises – a gap that was invisible in aggregated MSME data.

Maintain comparability
Use consistent indicator definitions, metadata, and measurement approaches across datasets and reporting periods. This ensures that results are comparable over time and across institutions, enabling accurate monitoring and benchmarking.

Example: The WE Finance Code's core indicators are defined consistently across participating countries, allowing performance comparisons and identification of regional trends (see Guidelines for the Collection and Reporting of Sex-Disaggregated Data on MSME Financing, WE Finance Code resources).

Validate and ensure methodological consistency
Maintain methodological consistency across sources and time periods so that aggregated results reflect genuine trends rather than changes in data coverage or methods. The data analysis and aggregation process should include systematic quality controls – such as outlier detection, internal consistency checks, and validation with alternative data sources – to ensure robustness. All aggregation or adjustment steps should be properly documented, including assumptions, estimation methods, and exclusion criteria, to promote transparency and replicability.

Maintain granularity but avoid unnecessary complexity
Disaggregated data should be detailed enough to reveal meaningful differences, but not so complex that results become unreadable or unmanageable. Grouping categories where appropriate can help maintain clarity without losing analytical value.

Example: The Dominican Republic's SIMBAR dashboard groups MSME loan data by micro, small, and medium categories, while still providing gender and sector filters for targeted analysis.

Provide context
Quantitative results should be interpreted in light of qualitative or contextual information, such as policy changes, market conditions, or sociocultural factors that may influence the numbers.

Example: Bangko Sentral ng Pilipinas pairs gender-disaggregated statistics with short MSME case profiles and photos to humanize the data and make policy implications clearer for non-technical audiences.

6.4 Dashboards for MSME data

Application programming interface (API) and dashboard integration for MSME data

Several countries have developed dashboards that aggregate and visualize MSME data in near real-time using APIs. These efforts build on a growing body of work from the Women's Financial Inclusion Data (WFID) Partnership, of which AFI is a member, and the Financial Alliance for Women, in collaboration with We-Fi and ConsumerCentriX. Together, these initiatives have supported regulators and financial institutions to collect, analyze, and visualize sex-disaggregated MSME and financial inclusion data.

Bangladesh – Central Bank MSME Dashboard[20]

The central bank's MSME dashboard connects via APIs to commercial bank reporting systems and the SME Foundation's datasets. This integration enables automatic monthly updates on indicators, such as MSME loan volumes, borrower counts, sectoral distribution, and geographic coverage. The dashboard is used internally for policy monitoring and externally to inform stakeholders of market trends.

Nigeria – WFID and SME Observatory Dashboards

Nigeria's SME Observatory, in partnership with FinTech companies, built an API-fed dashboard that combines traditional financial sector data (e.g. bank credit reports) with alternative data sources (e.g. mobile money transaction volumes). The platform disaggregates results by gender, business size, location, and sector, enabling targeted analysis of underserved groups, including informal MSMEs.

In parallel, the WFID Nigeria Dashboard (wfid.ng) was launched under the WFID Partnership, led by the Central Bank of Nigeria and supported by the Financial Alliance for Women and ConsumerCentriX. It compiles gender-disaggregated financial inclusion data and complements the SME Observatory's broader MSME indicators.

Note: Also see the Bangladesh WFID dashboard in 6.4.2 for a gender-specific platform with MSME-relevant indicators.

[20] This dashboard was developed with support from the WFID Partnership and the Financial Alliance for Women, with links to the global WE Finance Code framework.

Other good practices and platforms

Morocco – SME Observatory Centralized Dashboard (see Case Study 2 in Chapter 4)

The Morocco SME Observatory maintains a centralized dashboard drawing on tax, social security, credit bureau, and registry data. Gender identification is enhanced through machine learning applied to enterprise records. The platform enables multidimensional analysis by size, sector, location, revenue, employment, and value-added, providing policymakers and the financial sector with a robust, regularly updated MSME intelligence tool.

Solomon Islands – National Financial Inclusion Taskforce dashboard-style outputs (see Case Study 3 in Chapter 4)

The National Financial Inclusion Taskforce uses dashboard-style outputs from a new reporting template that collects gender-disaggregated indicators on deposits, loans, non-performing loans, and credit guarantee scheme participation. While primarily for internal use by the National Financial Inclusion Task Force, MSME Working group, and related stakeholders, the format enables quick policy insights and supports the country's move toward online portals to collect MSME data.

Bangladesh – Women's Financial Inclusion Data (WFID) Dashboard (link)

In 2024, Bangladesh Bank launched the WFID dashboard to monitor progress on women's access to and usage of financial services. The platform compiles gender-disaggregated data from banks, non-bank financial institutions, and mobile financial service providers, with indicators including account ownership, loan disbursements, savings balances, and repayment performance, all viewable by region and provider type. Although not MSME-specific, the dashboard tracks several indicators relevant to women-led MSMEs, providing a model for gender-responsive data transparency, and is publicly accessible via Bangladesh Bank's website.

WE Finance Code – Gender and MSME Data Platforms

The We Finance Code has developed interactive platforms to visualize and share gender-disaggregated MSME data across participating countries. These tools integrate data from multiple sources, including national statistical offices, financial sector regulators, business registries, and targeted surveys, into a single, filterable interface.

Users can explore indicators covering business demographics (e.g. size, sector, formality), access to finance (e.g. loan volumes, approval rates, NPLs), and usage of financial products, all disaggregated by gender. Filters allow cross-cutting views by country, region, sector, or financing type.

The platforms are designed for both **public access**, enabling policymakers, DFIs, and advocacy groups to benchmark progress, and **restricted partner access** for more granular, sensitive datasets. Country pages often combine tabular data with visual summaries, interactive maps, and trend charts.

In the Dominican Republic, the Superintendency of Banks adapted WE Finance Code's framework through the SIMBAR platform (see Case Study 5.1 in Chapter 4), consolidating credit and financial access indicators from regulated entities. Public-facing summaries and internal Power BI dashboards present MSME data progressively disaggregated by gender, informing both the annual MSME Report and the biannual Gender in Banking Report. This adaptation illustrates how an existing regulatory data system can integrate WE Finance Code's core indicators and gender perspective over time.

Additional tools and case materials from the **Financial Alliance for Women**, including Collecting and Using Banking Data on Women Businesses: A How-To Guide for Financial Institutions, Banco BHD's WMSME Data Journey, and Women's Markets Definitions (InBrief), offer practical guidance for financial institutions building gender-disaggregated data systems aligned with the WE Finance Code framework.

Lessons learned

Across both API-integrated MSME dashboards and other gender and MSME data platforms, several common lessons emerge:

1

AUTOMATE DATA FLOWS WHERE POSSIBLE

API connections between reporting systems reduce the manual reporting burden, improve timeliness, and free up staff time for analysis rather than data compilation.

Example: Bangladesh's MSME dashboard integrates directly with bank reporting systems and SME Foundation datasets, refreshing monthly without manual uploads.

2

INTEGRATE ACROSS DATA SILOS

Combining administrative sources (e.g. credit bureau, registry data) with alternative data (e.g. mobile money transactions, geospatial mapping) expands coverage, especially for informal and women-led MSMEs.

Example: Nigeria's SME Observatory merges bank credit reporting with mobile money transaction volumes to capture a broader picture of MSME activity.

3

PRIORITIZE GENDER-DISAGGREGATION

All platforms should enable filtering by gender, MSME size, location, and sector to make gender disparities visible and actionable.

Example: Morocco's SME Observatory applies gender identification (including AI-based inference) across its datasets to support multidimensional analysis.

4

PROVIDE REGULAR UPDATES

Dashboards with monthly or quarterly refresh cycles remain relevant for decision-making and avoid becoming static reports.

5

TAILOR ACCESS LEVELS TO AUDIENCE NEEDS

Provide public-facing summaries for transparency, while offering secure, detailed views to policymakers and technical users.

Example: The Dominican Republic's SIMBAR platform offers public summaries while internal dashboards provide more granular data for the Superintendency of Banks.

6

DESIGN FOR USABILITY AND CLARITY

Use clear visuals, intuitive filters, mobile-friendly layouts, and avoid overloading users with excessive indicators or complex charts.

Example: We-Fi's prototype platforms limit on-screen indicators to the most relevant for the selected filters, ensuring that even on mobile devices, users can quickly interpret gender gaps and MSME trends without scrolling through dense tables.

7

ADAPT PROVEN MODELS TO THE LOCAL CONTEXT

National platforms can adopt and customize frameworks such as We-Fi's or Morocco's Observatory model to fit their own systems, indicator sets, and policy priorities.

Example: Solomon Islands adapted its MSME indicators into dashboard-style outputs for the MSME Working Group and National Financial Inclusion Task Force.

6.5 Practical tips for communicating MSME and DDI data

Chart type recommendations
Choose chart types that match the story you want to tell. Use bar or column charts when comparing categories, such as MSME loan volumes by gender, line charts for showing trends over time, as well as maps for geographic patterns.

Example - Morocco: The SME Observatory applies geospatial mapping of MSME indicators (e.g. loan access, registry coverage) to highlight underserved regions, making spatial gaps visible at a glance.

Example - Solomon Islands: The Central Bank uses a simpler approach by downloading a display map from Google and manually inserting MSME credit access data when preparing quarterly reports. While not a full geospatial tool, this method still helps visualize regional access disparities with minimal resources.

Gender-sensitive color schemes
Use clear, consistent color-coding to distinguish gender without relying on stereotypical pink or blue palettes, which improves accessibility and avoids reinforcing gender clichés.

Example: **Bank Al-Maghrib** *and* **Solomon Islands CBSI** *apply carefully chosen, high-contrast color schemes for gender in charts, ensuring that differences are visible even for non-technical audiences.*

Annotations and storytelling in data presentation
Go beyond displaying raw numbers by guiding the audience through a narrative. Use annotations to explain significant changes, contextual notes to link results to real-world developments, and a logical flow from data to implication.

Example: **CGAP** *employs "data storyboards" in Power BI that lead viewers from problem > data > insight > policy implication, avoiding data dumps and reinforcing key messages with visual cues.*

7 CAPACITY BUILDING AND INSTITUTIONAL STRENGTHENING

Strengthening institutional capacity is essential for improving the quality, coverage, and use of MSME data. The DDI not only serves as a diagnostic tool to assess where an institution stands and identify priority areas for improvement, but also provides a roadmap for building the skills, systems, and governance arrangements needed to progress to higher levels of maturity. This chapter offers practical guidance on how institutions can:

- Diagnose their capacity needs by DDI maturity level and dimension.
- Follow tailored learning pathways suited to their role and mandate.
- Embed the toolkit and DDI approach into national data strategies, working groups, and technical assistance programs.

The goal is to help institutions move from assessment to action, ensuring that investments in MSME data systems are matched with the human, technical, and organizational capabilities to sustain them.

7.1 Capacity diagnostics by DDI level

Institutions at different stages of DDI maturity require different skills, resources, and systems to strengthen their MSME data capabilities. Because progress often varies across the four DDI dimensions (D1–D4), capacity diagnostics should be carried out **per dimension** rather than based solely on the overall score, so that targeted support effectively addresses the most pressing needs. In future phases, jurisdictions at more advanced maturity levels could also act as regional mentors or reference peers, voluntarily sharing their experiences in data integration and governance with other institutions across the AFI network.

Foundational level

Institutions at the Foundational stage are beginning to build the structures, skills, and tools needed for MSME data work. Priority needs often include:

- **D1 – Data collection**
  Establish basic coordination across departments; develop core Excel and data entry skills among staff; agree on working MSME definitions; pilot simple gender-disaggregated data collection.
- **D2 – Gender responsiveness**
  Raise awareness of the importance of gender-disaggregated data; add gender fields to existing forms and databases; introduce basic checks to ensure that gender fields are consistently completed.
- **D3 – Data usage and integration**
  Build capacity to produce simple reports using available data; develop skills in reading and interpreting basic charts and tables; initiate planning for future data integration.
- **D4 – Governance and infrastructure**
  Establish informal coordination mechanisms; designate focal points for MSME data; develop a first draft of standard operating procedures (SOPs); begin allocating a minimal budget for data activities.

PRACTICAL ROADMAP: BUILDING A GENDER-RESPONSIVE DATA SYSTEM (FIRST 12 MONTHS)

Institutions at the Foundational level can take a few simple, high-impact steps during their first year to embed gender responsiveness in their MSME data systems:

- Define and adopt a working definition for W-MSMEs.
- Add gender fields to MSME registration or enterprise survey forms.
- Map key stakeholders and assign focal points for gender data.
- Establish a basic data-sharing memorandum of understanding (MoU).
- Pilot API-based reporting with one or two financial service providers (FSPs).
- Produce and publish the first gender-disaggregated MSME data dashboard or report.

These early actions provide tangible momentum and lay the foundation for more advanced gender-responsive data practices as institutions progress across the DDI maturity levels.

Emerging level

At the Emerging stage, institutions have established some systems and are moving towards integration and more sophisticated use of MSME data. Key needs include:

- **D1 – Data collection**
  Harmonize definitions and indicator lists across agencies; maintain a registry of MSME indicators; train staff in integrating survey and administrative data sources.
- **D2 – Gender responsiveness**
  Strengthen skills for interpreting gender gaps; produce gender-responsive briefing notes; implement systematic quality control for gender-disaggregated fields.
- **D3 – Data usage and integration**
  Develop skills in dashboarding tools (e.g. Power BI, Tableau); link two or more core datasets; document indicator metadata for institutional use.
- **D4 – Governance and infrastructure**
  Formalize governance structures for MSME data; implement SOPs across agencies; secure dedicated budget lines; expand the network of designated focal points.

Established level

Institutions at the Established stage have robust systems and play a leadership role in MSME data work nationally or regionally, with a focus on sustaining excellence and expanding influence. Needs include:

- **D1 – Data collection**
  Conduct complex surveys and link multiple administrative and alternative datasets; produce regular, advanced disaggregations, including intersectional gender analysis.
- **D2 – Gender responsiveness**
  Institutionalize high-quality gender analysis in all outputs; embed gender data into policy cycles; mentor peer institutions on gender-responsive practices.
- **D3 – Data usage and integration**
  Implement advanced API-based integration for real-time dashboards; develop predictive analytics and scenario modelling; integrate dashboards into decision-making workflows.
- **D4 – Governance and infrastructure**
  Strengthen inter-agency data governance frameworks; secure sustainable funding; embed MSME data into policy feedback loops; actively participate in peer learning and south-south cooperation.

7.2 Learning pathways

Capacity building should be tailored to the specific roles and mandates of the different actors involved in MSME data systems. The three tracks below provide suggested focus areas for data producers, policymakers or regulators, and gender specialists. Institutions can adapt these to their own structures, resources, and DDI maturity levels.

Ethical and secure data management is a shared responsibility across all actors involved in MSME data systems. All institutions should ensure that staff are familiar with confidentiality and financial supervision regulations, and that disaggregated data is managed in ways that prevent the exposure of information about individual institutions or clients, in line with the principles outlined in Annex 4: Ethical and Confidentiality Principles for MSME Data Systems.

Data producers (e.g. NSOs, central bank data teams, IT units)

Focus on the technical and operational skills needed to ensure that MSME data is accurate, consistent, and interoperable.

- **Standards and definitions:** Apply harmonized MSME definitions and indicator frameworks across data sources.
- **Data processing and quality control:** Use validation rules, completeness checks (including for gender fields), and regular audits to improve data reliability.
- **Integration skills:** Learn to link survey, administrative, and alternative data sources; build capacity for API-based data exchange.
- **Metadata management:** Document sources, definitions, and methodologies for each indicator to ensure transparency and comparability.

Policymakers and regulators (e.g. ministries, financial regulators, SME agencies)

Focus on interpreting and applying MSME data to guide strategies, policy design, and monitoring.

- **Data interpretation:** Understand how to read and interpret DDI scores and MSME indicators, including dimension-level analysis.
- **Use cases:** Develop policy briefs, investment justifications, and monitoring reports grounded in disaggregated MSME data.
- **Strategic alignment:** Integrate MSME data into national development strategies, financial inclusion plans, and regulatory frameworks.
- **Feedback loops:** Establish mechanisms to provide feedback to data producers on data gaps and policy needs.

Gender specialists (within public institutions, regulators, or statistical offices)

Focus on embedding gender responsiveness into every stage of the MSME data cycle.

- **Disaggregation and analysis:** Ensure that all relevant indicators are disaggregated by gender, MSME size, and location, and interpret results to identify gaps.
- **Ethics and data protection:** Apply ethical standards to protect sensitive gender-disaggregated information, especially for vulnerable groups.
- **Advocacy and communication:** Use data to build internal buy-in for gender-responsive policies and resource allocation.
- **Capacity sharing:** Mentor colleagues in effectively understanding and using gender-disaggregated MSME data.

7.3 Toolkit adoption and country-level implementation

For the toolkit to have lasting impact, it needs to be embedded into existing institutional processes and national data ecosystems. Adoption should go beyond a one-time assessment to become part of the way institutions plan, monitor, and strengthen MSME data systems.

Options for embedding the toolkit

- **Internal self-assessment workshops**
  Organize workshops where cross-departmental teams score their institution against the DDI, identify gaps, and agree on priority actions, helping to build a shared understanding and ownership of the results.

- **Integration into national data strategies or working groups**
  Position the toolkit and DDI as part of national MSME data strategies, financial inclusion roadmaps, or sectoral working groups, promoting alignment with other data initiatives and facilitating coordination across agencies.

- **Use in technical assistance or south-south peer learning**
  Incorporate the toolkit into technical assistance projects, donor-funded programs, or peer-learning exchanges between countries, as comparing DDI profiles can help identify common challenges and successful approaches.

Suggested steps for institutionalization

1. **Appoint institutional focal points** for MSME data who are responsible for maintaining the DDI assessment and follow-up actions.
2. **Embed DDI indicators and targets into Key Performance Indicators** (KPIs) for relevant departments or agencies, ensuring accountability for progress.
3. **Integrate the DDI cycle into annual planning and review processes**, so that updates and capacity-building actions are tracked over time.
4. **Secure dedicated budget lines** for MSME data activities, linked to the capacity building priorities identified in the DDI.
5. **Establish feedback loops** between data producers, policymakers, and gender specialists to keep the toolkit relevant and responsive to evolving needs.

Suggested sequence for first-time use of the toolkit

1. **Preparation**
   - Identify the lead agency and focal points.
   - Gather existing MSME datasets, definitions, and documentation.
   - Brief senior management on the toolkit's goals and expected outputs.
2. **DDI self-assessment**
   - Convene a workshop with representatives from all relevant departments.
   - Score each DDI dimension based on current practices.
   - Document the evidence and rationale for each score.
3. **Gap analysis and action planning**
   - Use capacity diagnostics (7.1) to identify specific needs by dimension.
   - Map needs to learning pathways (7.2) for each user group.
   - Prioritize short-term "quick wins" and long-term system changes.
4. **Implementation (ongoing)**
   - Integrate actions into work plans and KPIs.
   - Schedule follow-up capacity building sessions.
   - Engage with peers or partner countries for experience sharing.
5. **Annual review and update**
   - Repeat the DDI assessment annually.
   - Track progress in each dimension and adjust plans accordingly.

CONCLUSION

Gender-responsive MSME data is more than a reporting requirement – it is a foundation for improved policy, smarter regulation, and more inclusive economic growth.

The journey from building data systems to implementing effective policies requires clear processes, sustained capacity, and a commitment to using evidence for change. At the same time, these data systems enhance prudential oversight by helping regulators and supervisors monitor risk concentration, credit trends, and the resilience of vulnerable market segments, linking financial inclusion with financial stability.

This toolkit offers a practical, modular pathway for institutions at all levels of maturity to strengthen their MSME data systems, embed gender responsiveness, and create a lasting impact. Drawing on the experience of AFI's SME Finance Working Group (SMEFWG) and Financial Inclusion Data and Impact Working Group (FIDIWG), as well as inputs from over 50 member institutions, it brings together proven approaches, capacity building guidance, and adaptable tools that work in diverse contexts.

When applied effectively, the toolkit enables institutions to:

- **Design and evaluate policies** that close gender gaps in MSME finance, from credit guarantee schemes that better reach women entrepreneurs, to procurement reforms that expand opportunities for women-led businesses.
- **Inform digital finance and innovation policies** that meet the needs of underserved MSMEs, including those in the informal sector.
- **Create feedback loops** where data directly shapes decision-making and policy outcomes feed back into improving data quality and coverage.

AFI's leadership and global network provide the platform for **shared learning and collective action**. By adopting this toolkit, member institutions can not only strengthen their own MSME data systems but also contribute to a broader movement for gender-inclusive finance.

Call to action:
Start by conducting a DDI self-assessment, map your capacity needs, and commit to integrating gender-responsive MSME data into your strategies, regulations, and reporting systems. Share your progress through AFI's working groups and engage in peer learning to further refine your approach.

ANNEXES

ANNEX 1.
SCORING LOGIC ON DATA DEVELOPMENT INDEX (DDI)

Data collection

What this dimension measures: Whether the institution collects micro-level MSME data, the **types of data** collected, the **presence of unique identifiers**, and the **methods used for processing** and ensuring reliability.

1.1 Micro-level data collection	Score	Descriptor	Rationale
No	0	Not collected	The institution does not collect any micro-level data on enterprises.
Yes, formal only	1	Partial coverage	The institution collects micro-level data only on formal enterprises (e.g. registered businesses in tax or credit registries).
Yes, formal + informal	2	Comprehen-sive	The institution collects micro-level data on both formal and informal enterprises. This includes those outside of official registration systems (e.g. via surveys, local directories, or alternative data sources).

1.2 Types of data collected	Score	Descriptor	Rationale
None	0	No data types	The institution does not collect or access any MSME-level data fields (e.g. no descriptive, identification, or financial data).
One type	1	Limited coverage	The institution collects only one category of MSME data (e.g. identifiers only or balance sheet data only).
Two or more types	2	Diverse coverage	The institution collects a combination of identification, descriptive, and financial data, enabling a more complete profile of MSMEs.

1.3 Use of unique identifiers	Score	Descriptor	Rationale
No unique identifier or cross-referencing system	0	No linkage	No consistent identifier is used; datasets are disconnected and cannot be reliably linked across systems. MSME datasets are siloed; no consistent identifiers exist to link across systems.
Use of multiple identifiers (e.g. tax, employment, business registry) but not unified	1	Fragmented linkage	Multiple identifiers exist and are used (e.g. tax ID, business registry, Social Security ID), but they are not unified. Institutions apply internal rules or ad hoc methods to cross-match records.
Use of a fully interoperable unique ID system across all major datasets	2	Integrated linkage	A single, interoperable identifier (e.g. national ID-linked business registration) is used consistently across data sources, supporting data integration and analysis.

1.4 Data processing methods	Score	Descriptor	Rationale
No structured method (e.g. ad hoc or manual with no SOPs)	0	Unstructured	No structured approach to data processing. Manual handling is ad hoc, without standard operating procedures (SOPs) or consistency across departments or data types. Data reliability and extraction are limited.
Structured manual or basic digital methods (e.g. Excel + SOPs or clear paper forms)	1	Basic processing	Data is processed through structured manual methods (e.g. standardized paper forms with SOPs) or basic digital tools (e.g. Excel templates, shared protocols). These methods allow for routine, replicable, and relatively reliable data use, even if not fully digitized.
Advanced statistical tools or automated systems (e.g. R, SAP, API-fed dashboards)	2	Advanced processing	Advanced tools and systems are in place, including statistical software (e.g. R, SPSS), automated platforms (e.g. SAP), or API-enabled dashboards. These enable integrated, scalable, and timely analysis with reduced human error.

Aggregation rule
Raw total: **0 to 7**
Scoring:
0-1 > 0
2-3 > 1
4-5 > 2
6-7 > 3

0 SCORE 0 – NO STRUCTURED COLLECTION SYSTEM

- The institution does not collect MSME-level data.
- No access to microdata from registries, surveys, or partner institutions.
- Data is handled manually or on demand, with no SOPs or validation protocols.
- Datasets (if any) cannot be linked or reused due to the lack of unique identifiers.

Example: A ministry that only receives high-level MSME statistics from the national statistical office once a year and keeps internal records informally in PDFs or emails.

1 SCORE 1 – EARLY, FRAGMENTED COLLECTION

- The institution collects some MSME-level data, but it is limited to formal enterprises.
- Data comes from a single administrative source (e.g. tax registry or credit bureau).
- One type of data is collected (e.g. ownership or sector only).
- Processing is done using Excel or basic systems, guided by internal SOPs.
- Multiple IDs may exist, but they are not linked across systems.

Example: A central bank that collects balance sheet data from banks but does not include informal MSMEs or non-financial information.

2 SCORE 2 – STRUCTURED COLLECTION AND PARTIAL INTEGRATION

- The institution collects data on both formal and informal MSMEs (e.g. through surveys or partnerships).
- It captures multiple types of data (e.g. identifiers, performance, access to finance).
- Uses multiple identifiers (e.g. tax ID + business number), though not fully interoperable.
- Data processing includes SOPs and basic tools (e.g. R, SPSS) for analysis and reporting.

Example: A national statistics office integrates data from business surveys and tax records, processed with scripts in SPSS and partially linked using business IDs.

3 SCORE 3 – ADVANCED, INTEGRATED COLLECTION SYSTEM

- The institution collects comprehensive MSME data, covering formal and informal sectors, multiple types, and key demographic fields.
- Data is linked across systems via interoperable unique identifiers.
- Processing is automated using enterprise platforms, statistical scripts, or API-linked dashboards.
- Data systems are well-documented, regularly updated, and support policy design.

Example: An institution like India's Udyam Portal or Morocco's SME Observatory, where data is collected centrally, linked across systems, processed via APIs, and used for real-time dashboards.

Gender responsiveness

What this dimension measures: Whether institutions collect and use gender-disaggregated data to understand W-MSMEs, including ownership, leadership, access to credit, DFS use, and broader gender-specific indicators.

2.1 Collection of the owner or manager's gender	Score	Descriptor	Rationale
Not collected	0	Not collected	The institution does not collect or have access to information on the gender of the MSME owner or manager, and no disaggregation is possible.
Collected or accessible	1	Available	The gender of the MSME owner or manager is collected (e.g. via administrative datasets, registration forms, or surveys) or can be systematically inferred through a documented method such as name-matching or machine learning. The data is usable for gender disaggregation and analysis.

2.2 Collection of gender-disaggregated or gender-relevant indicators	Score	Descriptor	Rationale This criteria assesses whether the institution collects indicators that are disaggregated by gender (e.g. access, usage, outcomes) or specifically relevant to gender analysis (e.g. women's leadership or training participation). It goes beyond simply recording the gender of the owner or manager to include data that reveal gender gaps and inform responsive policymaking.
No indicators collected	0	Not collected	The institution does not collect any gender-disaggregated or gender-relevant MSME indicators – beyond the gender of the owner or manager. There is no data on women's leadership, access to finance, use of services, etc.
One indicator collected	1	Partial coverage	The institution collects one gender-disaggregated or gender-relevant MSME indicator. This may include the share of women-owned or women-led enterprises, loan access by gender, participation in financial training by gender, etc.
Two or more indicators collected	2	Expanded coverage	The institution collects two or more such indicators relevant to W-MSMEs. These may include leadership (e.g. CEO, board participation), credit access and terms (loan amounts, approvals, interest rates), DFS usage (mobile money, e-wallets), training, financial literacy, or formalization by gender.

2.3 Disaggregation of DFS usage by gender	Score	Descriptor	Rationale
Not disaggregated	0	Not disaggregated	The institution does not collect any data on DFS usage among MSMEs, or it does collect such data but does not disaggregate it by gender. As a result, it cannot assess gender gaps in digital finance access or use.
Disaggregated by sex	1	Disaggregated by sex	The institution collects data on DFS usage and disaggregates it by gender. This enables analysis of differential access or behavior in DFS among women- and men-led MSMEs. Innovative proxy methods (e.g. matching DFS user records to gender-tagged business data or using predictive models) may also qualify if applied systematically and transparently.

Aggregation rule

Raw total: **0 to 4**

Scoring:

0 > 0
1 > 1
2-3 > 2
4 > 3

0 SCORE 0 – NO GENDER DATA PRACTICES IN PLACE

- The institution does not collect gender data for MSME owners or managers.
- No indicators are disaggregated by gender.
- No data on women's access to DFS or credit exists, and no effort is made to fill these gaps.
- Gender analysis is not possible due to the absence of structured data.

Example: A financial regulator tracks MSME portfolios from banks, but none of the data is disaggregated by gender – and ownership or leadership information is absent from registration forms.

1 SCORE 1 – INITIAL OR PARTIAL GENDER RESPONSIVENESS

- The gender of MSME owners or managers is collected or inferred (e.g. via survey or business registry).
- The institution collects one gender-specific indicator (e.g. ownership or loan approval by sex).
- DFS usage data is collected but not yet disaggregated, or only partial attempts are made.
- Gender is considered but not institutionalized.

Example: A ministry collects survey data that includes ownership by gender but does not include other gender indicators, and while DFS data is available, it is not analyzed by sex.

2 SCORE 2 – BROADER BUT STILL LIMITED GENDER INTEGRATION

- Gender is systematically captured in enterprise records (e.g. registration or credit data).
- The institution collects two or more gender-specific indicators (e.g. leadership, DFS use, access to credit).
- Data is disaggregated by sex for at least some program or policy planning.
- Gender is considered in analysis but not yet used in real-time or dashboard formats.

Example: A central bank collects data from FSPs on loan approval rates and DFS use by sex and includes this information in quarterly internal reports.

3 SCORE 3 – FULL INTEGRATION OF GENDER ACROSS THE MSME DATA SYSTEM

- Gender fields are embedded in enterprise registration, credit reporting, or survey systems.
- The institution collects a broad set of gender-specific indicators, disaggregated and updated regularly.
- DFS data is gender-disaggregated and used to inform program design, monitoring, or dashboards.
- Gender-disaggregated data is routinely used in decision-making and reporting.

Example: An SME Observatory publishes quarterly dashboards on W-MSME performance, using sex-disaggregated data on credit, DFS usage, leadership, and training participation.

Data usage and integration

What this dimension measures:
Whether MSME data (especially gender-disaggregated) is used for policy, program design, or strategic planning, and whether systems are integrated, aligned with global standards, and supported by institutional awareness.

3.1 Level of data integration across sources	Score	Descriptor	Rationale
No integration of data	0	Siloed data	Data is collected and stored by individual departments or systems without any integration. Datasets (e.g. tax, finance, gender) remain separate and cannot be cross-analyzed.
Integration within the same source only	1	Partial integration	Data is integrated within the same system or dataset (e.g. from different branches of a registry or financial system), but is not linked across institutions or sectors.
Integration from multiple sources	2	Cross-source linkage	MSME data is integrated across multiple systems or sources (e.g. tax records + financial data + social security), enabling comprehensive analysis and policymaking. Data-sharing agreements, APIs, or interoperable platforms support this.

3.2 Alignment with international standards	Score	Descriptor	Rationale
Not aligned	0	No alignment	The institution does not reference or apply international frameworks (e.g. Global Findex, IMF FAS, World Bank Enterprise Surveys) when defining indicators, variables, or reporting formats.
Partially aligned	1	Partial alignment	The institution uses some elements of international standards (e.g. indicator definitions or formats) but without full consistency. Some key variables (e.g. ownership, leadership, access to credit) may differ from global benchmarks.
Fully aligned	2	Full alignment	Indicator definitions and reporting formats follow recognized international frameworks. This enables comparability and strengthens interoperability with global datasets or donor systems.

3.3 Awareness and use of gender data tools/resources	Score	Descriptor	Rationale
No	0	No awareness	The institution is not aware of available tools, frameworks, or initiatives focused on gender-disaggregated data (e.g. Data2X, WFID, AFI gender guidelines). Staff are not using reference materials or participating in knowledge-sharing activities on gender and data.
Yes	1	Aware and engaged	The institution is aware of external gender data tools and resources, such as those developed by AFI, Data2X, WFID, or ITC SheTrades, and has engaged with them in some way. This may include participation in training sessions, internal referencing of guidance materials, or informal use of templates or case studies. This score reflects institutional openness to improving gender data practices even if full integration into systems or policy has not yet occurred.

Aggregation rule
Raw total: **0 to 5**
Scoring:
0 > 0
1-2 > 1
3-4 > 2
5 > 3

0 SCORE 0 – NO DATA USAGE OR INTEGRATION

- MSME data is collected but not used in decision-making or policy documents.
- Datasets are siloed with no integration across sources.
- No awareness of international standards or gender data tools.
- Data remains internal, fragmented, and underused.

Example: A statistics unit collects enterprise survey data but does not share it with policymakers or publish gender findings, while departments work independently with no shared templates.

1 SCORE 1 – BASIC USAGE OR LIMITED INTEGRATION

- MSME data is referenced occasionally in reports or assessments.
- Data is integrated within a single source, such as one ministry or agency.
- Partial alignment with international standards.
- Staff are aware of gender data tools or attended at least one training.

Example: A central bank aligns some credit indicators with FAS definitions and is exploring SheTrades dashboards, but gender data is not consistently used.

2 SCORE 2 – SYSTEMATIC USE AND CROSS-SOURCE INTEGRATION

- Gender is systematically captured in enterprise records (e.g. registration or credit data).
- The institution collects two or more gender-specific indicators (e.g. leadership, DFS use, access to credit).
- Data is disaggregated by sex for at least some program or policy planning.
- Gender is considered in analysis but not yet used in real-time or dashboard formats.

Example: A financial regulator builds a quarterly policy brief using multi-source data and tracks gender gaps in loan approvals, based on standardized templates.

3 SCORE 3 – FULL DATA INTEGRATION AND STRATEGIC USE

- Gender-disaggregated MSME data is used routinely to inform policy, budget decisions, and program design.
- Data systems are interoperable and updated in real time.
- Full alignment with global frameworks.
- The institution participates in regional learning, produces open dashboards, and actively mentors others.

Example: A central bank integrates tax, DFS, and registry data into a live dashboard aligned with WFID and Findex. The data is used to adjust SME support schemes quarterly.

Governance and infrastructure

What this dimension measures:
Whether institutions have the internal systems, coordination structures, and resources to manage MSME data collection, integration, and usage in a sustainable way.

4.1 Existence and use of a unified platform for MSME data	Score	Descriptor	Rationale
No platform or access	0	Not connected	The institution does not access or participate in any unified or federated MSME data platform. Data is siloed across institutions or internal departments.
Institution participates in a platform managed by another agency (with access only)	1	Connected	The institution contributes to or uses a unified platform managed by another agency, such as a national observatory or registry. However, it does not lead or coordinate the platform.
Institution coordinates or co-coordinates a multi-source MSME data platform (ownership not required)	2	Cross-source linkage	The institution coordinates or co-coordinates a functioning MSME data platform that draws from multiple sources. It plays an active role in ensuring data sharing, updates, and policy relevance. Platform ownership is not required – coordination and functionality are the key criteria.

4.2 Existence of a dedicated unit for MSME data collection and processing	Score	Descriptor	Rationale
No dedicated unit	0	Not established	The institution does not have a unit or team specifically responsible for MSME or W-MSME data collection, processing, or analysis. Related responsibilities are either absent or dispersed across departments without coordination.
Dedicated unit exists	1	Functional	There is a clearly identified unit, team, or department responsible for MSME and gender-disaggregated data. This unit is active and has defined roles, even if small in size or limited in scope.

4.3 Perception of institutional resources for MSME data work	Score	Descriptor	Rationale
Disagree or strongly disagree	0	Inadequate resources	The institution does not have sufficient financial, technical, or human resources to support MSME data collection and analysis. Staff may lack time, systems may be outdated, or budgets may be missing.
Neutral	1	Limited resources	Some resources are available, but capacity is constrained. The institution may rely on external support, face staff or system limitations, or only engage intermittently in MSME data work.
Agree or strongly agree	2	Adequate resources	The institution perceives itself as having sufficient capacity, including funding, trained personnel, and basic infrastructure, to effectively collect, process, and analyze MSME and W-MSME data.

Aggregation rule
Raw total: **0 to 5**
Scoring:
0 > 0
1-2 > 1
3-4 > 2
5 > 3

0 SCORE 0 – NO INSTITUTIONAL STRUCTURES IN PLACE

- The institution does not participate in any unified MSME data platform.
- There is no dedicated unit or staff for MSME data work.
- Staff report insufficient time, tools, or resources to manage MSME or gender-disaggregated data.
- No defined roles, responsibilities, or coordination mechanisms exist.

Example: A ministry has no internal data team and only accesses MSME data ad hoc. There is no platform, no integration with others, and staff report budget and IT constraints.

1 SCORE 1 – BASIC STRUCTURES OR PARTIAL ENGAGEMENT

- The institution has access to a unified platform managed by another agency but does not coordinate it.
- There is a small team or focal point for MSME data, but coordination is informal.
- Resources are limited – tools may be outdated, or staff are juggling multiple roles.

Example: A financial regulator contributes data to a national observatory but has no internal coordination structure. One economist handles MSME data as a side task.

2 SCORE 2 – DEFINED SYSTEMS AND EMERGING COORDINATION

- The institution plays a coordinating role in a multi-source MSME platform or observatory.
- It has a dedicated unit for data collection and processing.
- Perceived resource levels are sufficient for routine MSME data work, with support in place.

Example: A central bank leads a shared MSME platform with the national statistics office and tax agency. A three-person unit manages MSME dashboards and analytics.

3 SCORE 3 – INSTITUTIONALIZED GOVERNANCE AND INVESTMENT

- MSME data systems are fully embedded in institutional structures and supported by strong internal mandates.
- The institution leads or co-leads an interoperable platform with defined data-sharing agreements.
- A well-resourced unit with clear roles, secure systems, and cross-departmental collaboration manages data flows and analysis.

Example: A national SME observatory under the central bank coordinates a platform integrating administrative and survey data, with staff dedicated to analytics, partnerships, and dashboard publication.

Final DDI aggregation

Final DDI Score	Condition	What it suggests
Early stage	Sum of dimension scores = 0 to 3	Basic or fragmented MSME data practices. Systems may be ad hoc, siloed, or not yet operational.
Emerging	Sum of dimension scores = 4 to 7	Core structures are in place, but coordination, integration, or gender disaggregation remain partial or inconsistent.
Advanced	Sum of dimension scores = 8 to 12	Strong systems and coordination across dimensions. MSME data is used strategically, with embedded gender responsiveness and institutional support.

Interpreting DDI results

The DDI is designed as a **diagnostic and developmental tool**, not a benchmarking system. While the final score gives a high-level indication of institutional maturity, its greatest value lies in examining **each of the four dimensions separately**.

Institutions may score well in one area, such as data usage, but face challenges in another, such as gender responsiveness. The same overall score can reflect **very different realities**, depending on the strengths and gaps across dimensions.

Use the DDI to identify your institution's **entry point for progress** – not to compare it with others.

ANNEX 2. DATA DEVELOPMENT INDEX (DDI) SURVEY

Dimension 1: Data Collection

1.1 Does your institution collect micro-level data on MSMEs?

- ○ No
- ○ Yes, but only for formal enterprises
- ○ Yes, for both formal and informal enterprises

1.2 What types of data does your institution collect on MSMEs?
(Select all that apply)

- ○ Identification data (e.g. business ID, location, legal forms)
- ○ Descriptive data (e.g. revenue, employment, export activity)
- ○ Financial data (e.g. loan approvals, credit terms)
- ○ None of the above

1.3 How are MSMEs identified across datasets in your institution?

- ○ No consistent identifiers or linkage methods
- ○ Multiple identifiers used but not interoperable (e.g. tax ID, registry number)
- ○ A single or interoperable identifier used across major datasets

1.4 How is your MSME data processed?

- ○ Manually or informally, with no SOPs or structure
- ○ Using structured manual methods or basic tools (e.g. Excel with SOPs)
- ○ Using automated tools or statistical platforms (e.g. APIs, R, SAP)

Dimension 2: Gender Responsiveness

2.1 Is the gender of MSME owners or managers collected or available in your data?

- ○ No
- ○ Yes. Directly collected or inferred through documented methods (e.g. name-based prediction)

2.2 Which gender-specific indicators does your institution collect (if any)?
(Select all that apply)

- ○ Women's ownership share
- ○ Women in leadership (e.g. CEO, board)
- ○ Gender-disaggregated access to finance (e.g. loans)
- ○ Use of DFS (e.g. mobile money) by gender
- ○ Participation in training (by gender)
- ○ Informal vs. formal employment (by gender)
- ○ None of the above

2.3 Does your institution disaggregate DFS usage data by gender?

- ○ No
- ○ Yes

Dimension 3: Data Usage and Integration

3.1 Are MSME data sources integrated in your institution?

- ○ No, data is siloed and unlinked
- ○ Yes, within the same department or system only
- ○ Yes, across multiple data sources (e.g. tax, credit, registries)

3.2 Does your institution align its MSME indicators with international standards?
(e.g. Global Findex, IMF FAS, World Bank definitions)

- ○ No alignment
- ○ Partial alignmen
- ○ Full alignment

3.3 Is your team aware of and using any external gender data tools or frameworks?
(e.g. AFI guidelines, WFID, SheTrades, Data2X)

- ○ No
- ○ Yes, we have used or referenced such tools or participated in training sessions

Dimension 4: Governance and Infrastructure

4.1 Does your institution participate in a unified platform for MSME data?

- ○ No
- ○ Yes, we access or contribute to a platform managed by another institution
- ○ Yes, we coordinate or co-coordinate a multi-source MSME data platform

4.2 Is there a dedicated unit or team in your institution responsible for MSME data collection and processing?

- ○ No
- ○ Yes

4.3 Do you agree that your institution has adequate resources (staff, budget, tools) to manage MSME and W-MSME data?

- ○ Strongly disagree
- ○ Disagree
- ○ Neutral
- ○ Agree
- ○ Strongly agree

ANNEX 3. GLOSSARY OF INDICATORS

Domain	Indicator	Tier	Description
Ownership and leadership	% ownership by gender	Tier 1	Proportion of business ownership held by women, men, or jointly. Should be disaggregated by gender to assess women's participation in MSME ownership.
	Gender of manager or CEO	Tier 1	Gender of the principal decision-maker in the MSME, such as the manager, CEO, or lead operator. Important to capture leadership roles beyond legal ownership.
	Age of owner	Tier 2	Age of the primary owner of the MSME. Can be disaggregated by gender to analyze age-gender dynamics in entrepreneurship.
	Education level of owner	Tier 2	Highest level of education attained by the owner. Useful for understanding the human capital behind MSMEs.
Business profile	MSME size classification (micro, small, medium)	Tier 1	Standard classification of MSMEs based on employee count, revenue, or assets. Typically used to structure financial and regulatory programs.
	Business registration or formality status	Tier 1	Whether the business is formally registered with authorities. Key for understanding informality and access to services.
	Business sector (e.g. trade, agriculture)	Tier 1	Economic sector in which the MSME operates. Important for identifying sectoral gender patterns or financing needs.
	Geographic location (region, rural or urban)	Tier 1	Location of the MSME, classified by region or urban/rural split. Crucial for spatial equity and inclusion analysis.
	Age of business	Tier 2	Number of years since the MSME was established. Helps assess business maturity and survival trends.
Access to finance	Number of MSME borrowers (WE Finance Code core indicator 1)	Tier 1	Total number of MSMEs with at least one active loan, disaggregated by gender of owner/leader and MSME size category.
	Volume of financing to MSMEs (WE Finance Code core indicator 2)	Tier 1	Total outstanding or disbursed value of loans to MSMEs during the reporting period, disaggregated by gender of owner/leader and MSME size category.
	Level of deposits for MSMEs (WE Finance Code core indicator 3)	Tier 1	Total value of deposits held in accounts by MSMEs, disaggregated by gender of owner/leader and MSME size category.
	Non-performing loans (NPLs) for MSMEs (WE Finance Code core indicator 4)	Tier 1	Value or percentage of MSME loans classified as non-performing according to the national regulatory definition, disaggregated by gender of owner/leader and MSME size category.

Access to finance	Loan approval rates for MSME applications (WE Finance Code core indicator 5)	Tier 1	Percentage of MSME loan applications approved during a given period, calculated as number of approvals divided by number of applications, disaggregated by gender of owner/leader and MSME size category.
	Loan application status	Tier 1	Tracks whether MSMEs apply for loans and whether applications are approved or denied. Should be disaggregated by gender to assess credit access.
	Loan terms (interest rate, maturity, collateral)	Tier 1	Conditions attached to loans received by MSMEs, including interest rates, repayment periods, and required collateral.
	Use of savings or deposit products	Tier 2	Whether the MSME uses savings accounts, fixed deposits, or similar products. Indicates financial inclusion beyond credit.
	Use of alternative finance (e.g. microfinance, FinTech)	Tier 2	Use of non-traditional finance sources such as microfinance institutions, peer-to-peer platforms, or FinTech services.
	Credit bureau presence or credit history	Tier 2	Whether the MSME or owner has a formal credit history recorded in a bureau or registry. Reflects financial visibility and eligibility.
	Number of financial accounts or services used	Tier 2	Total number of distinct financial services or accounts held (e.g. loans, insurance, mobile money). Reflects depth of financial engagement.
Digital access	Use of DFS (e.g. mobile money, online platforms)	Tier 1	Use of digital financial services such as mobile wallets, online banking, or digital credit tools.
	Access to internet or smart-phone for business purposes	Tier 2	Whether the MSME uses internet access or smartphones for business operations. Reflects digital enablement.
Public procure-ment	Participation in public contracts or tenders	Tier 2	Whether the MSME has applied for or received public sector contracts. Important for inclusion in public procurement systems.
	Value or % of contracts awarded	Tier 2	Monetary value or percentage of contracts awarded to the MSME. Useful for assessing scale of participation in procurement.
Business outcomes	Revenue or sales trends	Tier 2	Total revenue or change in sales over time. Core measure of MSME performance and growth.
	Use of DFS (e.g. mobile money, online platforms)	Tier 2	Whether the business has closed or remains active over time. Can be linked to economic resilience and barriers.
	Access to internet or smart-phone for business purposes	Tier 2	Number of jobs created by the MSME, ideally disaggregated by gender. Captures the contribution of MSMEs to employment and gender equity.

ANNEX 4. ETHICAL AND CONFIDENTIALITY PRINCIPLES FOR MSME DATA SYSTEMS

1. Purpose and Scope

This annex provides guiding principles for the ethical collection, processing, sharing, and use of MSME data, including gender-disaggregated information.

It aims to ensure that all data activities under the toolkit respect privacy, confidentiality, and human rights, while promoting transparency and accountability in accordance with international best practices.

2. Core ethical principles

Principle	Description	Illustrative application
1. Consent and transparency	Data subjects should be informed of the purpose and use of the data and, where applicable, consent should be obtained prior to collection or sharing.	Include consent clauses in business surveys; publish summaries of how MSME data is used in policymaking.
2. Purpose limitation and proportionality	Data should be collected and used only for clearly defined, legitimate purposes and limited to what is necessary for those purposes.	Restrict data use to policy design, monitoring, and research; avoid secondary use without authorization.
3. Minimization of identification risk	Data should be anonymized or aggregated to prevent the reidentification of individuals or specific enterprises.	Apply suppression or noise techniques when publishing regional or sectoral data.
4. Statistical-use only	MSME data, particularly from financial institutions, should be used strictly for analytical and policy purposes, not for enforcement or commercial gain.	Establish legal clauses or MoUs restricting use to statistical and supervisory analysis.
5. Data access traceability	All access, sharing, and modification of data should be logged to ensure accountability and enable audits.	Maintain access logs for users of MSME databases or dashboards.
6. Non-discrimination and fairness	Data analysis should avoid reinforcing gender or sector biases and ensure fair representation of informal and small enterprises.	Use balanced sampling methods; include validation to prevent systemic bias.
7. Accountability and oversight	Institutions managing MSME data should have designated officers responsible for compliance with ethical and confidentiality standards.	Appoint a Data Protection Focal Point or Ethics Officer in the MSME Data System.

3. Institutional arrangements

AFI members are encouraged to:

- Establish internal data ethics committees or focal points within statistical or supervisory departments.
- Adopt data-sharing agreements (MoUs) that explicitly include ethical clauses and confidentiality provisions.
- Conduct regular audits and training on responsible data management practices.

Action	Responsible entity	Frequency
Designate a Data Ethics Focal Point	Central bank, Statistical Agency, Data Ethics Focal Point, Data Protection Authority	Once – reviewed annually
Maintain data access and sharing logs	Data management unit	Continuous
Review data-sharing MoUs for ethical compliance	Legal, compliance units	Annual
Conduct staff training on ethical data handling	HR, capacity building department	Semi-annual
Publish transparency reports on MSME data use	Lead institution	Annual

ACRONYMS

AFI	Alliance for Financial Inclusion
API	Application Programming Interface
CGAP	Consultative Group to Assist the Poor
DDI	Data Development Index
FIDIWG	Financial Inclusion Data and Impact Working Group
FSP	Financial Service Providers
GDPR	European Union General Data Protection Regulation
MFI	Micro Finance Institutions
MNO	Mobile Network Operators
MoU	Memorandum of Understanding
MSMEs	Micro, Small and Medium Enterprises
NGO	Non-governmental Organizations
NSO	National Statistics Offices
SMEFWG	Small and Medium Enterprises Finance Working Group
WFID	Data2X's Women's Financial Inclusion Data partnership
WMSMEs	Women-Micro, Small and Medium Enterprises (refers to women-led and women-owned enterprises)
WEFi Code	Women Entrepreneurs Finance code (World Bank initiative)

Alliance for Financial Inclusion
AFI, Sasana Kijang, 2, Jalan Dato' Onn, 50480 Kuala Lumpur, Malaysia
t +60 3 2776 9000 **e** info@afi-global.org www.afi-global.org
Alliance for Financial Inclusion AFI.History @NewsAFI @afinetwork