



MINISTRY OF ECONOMY AND
FINANCE OF THE
REPUBLIC OF UZBEKISTAN



Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Agency for Development and Cooperation SDC
Shveysariya taraqqiyot va hamkorlik Agentligi SDC



MONITORING AND EVALUATION GUIDE

The views expressed in this report do not necessarily reflect the views of the United Nations, including UNDP, or the UN Member States. All reasonable precautions have been taken to verify the information contained in this publication. However, published material is distributed without any warranty, express or implied. UNDP is not responsible for the content of any external website. The designations and terminology used may not be in accordance with United Nations practice and do not imply the expression of any opinion by the Organization.

The study is conducted within the framework of the joint project of the Ministry of Economy and Finance of the Republic of Uzbekistan and the United Nations Development Programme (UNDP) "Empowering the Youth to Embrace the Digital Economy and Digital Entrepreneurship," financed by the Swiss Agency for Development and Cooperation (SDC).

The United Nations Development Programme (UNDP) is the leading UN organization fighting injustice caused by poverty, inequality and climate change. Working with a wide network of experts and partners in 170 countries and territories, we help to create integrated, long-term solutions for people and the planet.

Find out more about us at undp.org/uzbekistan or follow @UNDPUzbekistan on social networks. © UNDP 2025.

All rights reserved. The contents and information of this report may be reproduced elsewhere in whole and/or in part, provided the source is acknowledged. Source citation: Digital economy of Uzbekistan: The State of Digital Entrepreneurship to Harness AI, UNDP, 2025.

CONTENTS

1. INTRODUCTION.....	4
2. BACKGROUND	6
2.1. BASIC CONCEPTS	6
2.2. PRE-REQUISITES FOR A SUCCESSFUL M&E SYSTEM.....	7
2.2.1. POLICY PLANNING AS AN ENABLER FOR M&E: RESULTS BASED MANAGEMENT	8
2.2.2. DEVELOPMENT OF MONITORING PLAN.....	10
2.2.3. DEVELOPING OF EVALUATION PLAN.....	13
3. EVALUATION TYPOLOGY	16
4. EVALUATION METHODS.....	19
4.1. TOOLS FOR EVALUATION	25
4.2. PROCESS FOR SELECTING EVALUATION TYPE	30
4.2.1. STEP 1: IDENTIFY THE EVALUATION TRIGGER	30
4.2.2. STEP 2: DEFINE WHAT IS BEING EVALUATED	30
4.2.3. STEP 3: DETERMINE THE TIMING IN THE LIFECYCLE	31
4.2.4. STEP 4: CLARIFY THE PURPOSE OR INTENDED USE.....	31
4.2.5. STEP 5: SELECT THE EVALUATION TYPE.....	32
4.2.6. STEP 6: CONFIRM TOOLS, STAKEHOLDERS, AND READINESS.....	33
5. MONITORING IN A NATIONAL M&E SYSTEM	34
5.1. TYPES AND PURPOSES OF M&E	34
5.2. MONITORING ACROSS THE RESULT CHAIN	35
5.3. MONITORING FUNCTIONS AND RESPONSIBILITIES.....	36
5.4. MONITORING TOOLS AND MODALITIES	37
6. A CONCISE GUIDE TO REGULATORY IMPACT ASSESSMENT (EX-POST RIA)	39
6.1. WHAT IS EX-POST RIA AND WHY IT MATTERS	39
6.2. A STEP-BY-STEP GUIDE TO CONDUCTING EX-POST RIA	40
7. A CONCISE GUIDE TO POLICY IMPACT ASSESSMENT.....	47
7.1. INTRODUCTION TO EX-POST POLICY IMPACT ASSESSMENT (PIA)	47
7.1.1. WHY A POLICY-CENTRIC EVALUATION APPROACH?.....	47
7.1.2. THE LOGIC OF EX-POST POLICY IMPACT ASSESSMENT (PIA)	48
7.1.3. CONTRIBUTION OVER ATTRIBUTION.....	49
7.2. STEP-BY-STEP GUIDE TO CONDUCTING AN EX-POST PIA.....	50
7.2.1. USING THE STARTUP ECOSYSTEM SUPPORT POLICY AS AN ILLUSTRATIVE EXAMPLE	50
8. CONCLUDING REMARKS	55

1. INTRODUCTION

This Methodology Guide for Monitoring and Evaluation (M&E) has been developed as part of the technical assistance provided under the UNDP project: "*Enhancing Capacities of Public Servants in Developing, Monitoring and Evaluating Digital Entrepreneurship and Digital Economy Policies*". The Guide complements a draft **M&E Decree** being prepared in parallel under the same project which will define the formal roles, responsibilities, and coordination structures governing M&E across the public administration.

While the Decree will provide the legal and procedural scaffolding of a national M&E system, this Guide serves as its **practical counterpart**. It offers step-by-step methodologies, planning tools, and templates to help ministries, agencies, and oversight institutions operationalize the provisions of the Decree. It is designed to translate high-level intent into working processes, supporting the integration of M&E across policy formulation, regulatory practice, program implementation, and strategic planning.

The Guide builds on international good practices drawing on the OECD DAC evaluation framework and UNDP RBM guidance, while being carefully adapted to Uzbekistan's current institutional context. Alongside the Decree, it forms part of a broader reform package that also includes institutional diagnostics, capacity-building, and the establishment of performance-linked policy and regulatory processes.

The methodology and structure of the Guide reflect the project's broader aims: to strengthen the use of evidence in policy and regulatory decision-making; to embed performance-based thinking into government operations; to foster strategic alignment between ministries and center-of-government bodies; and to enhance technical capacity for designing, implementing, and using monitoring and evaluation practices across the public sector.

This Guide is intended to serve both as a **technical reference document** and as a **practical handbook** for those involved in designing or overseeing public interventions whether policies, programs, or regulations. It is structured to support users at different levels of government, from senior decision-makers and strategic planners to sector specialists, analysts, and M&E practitioners. The document moves from general to specific, starting with foundational concepts and moving toward applied techniques.

It begins with an overview of the core concepts that underpin results-based management and M&E in the public sector, clarifying the difference between outputs, outcomes, and impacts, and explaining the importance of intervention logic and indicator systems. This is followed by a chapter on the **preconditions for a functioning M&E system**, including the development of results frameworks, baselines, and target setting. The next section outlines the strategic planning functions that must be led or coordinated by the **center of government** particularly in relation to rolling evaluation planning, monitoring standardization, and alignment with national strategies.

Building on this foundation, the Guide then moves to hands-on guidance for public officials responsible for conducting or managing evaluations. This includes an in-depth typology of evaluation types, a process for selecting the appropriate evaluation method, and two dedicated chapters for applying **Policy Impact Assessment (PIA)** and **Regulatory Impact Assessment (RIA)** both focused on ex-post evaluation. These sections are supported by practical tools to assist in evaluation scoping, methodology selection, data collection, and reporting.

The Guide concludes with a modular set of **tools and templates** that can be adapted across ministries and sectors. These include sample monitoring frameworks, theory of change diagrams, indicator matrices, scorecards, evaluation question templates, and standard terms of reference. Each chapter is designed to stand alone or be read in sequence, with cross-references and examples throughout to support institutional learning and practical application.

In combination with the legal and institutional framework to be introduced via the Decree, this Guide provides a concrete foundation for building a credible, coordinated, and learning-oriented M&E system in Uzbekistan one that supports performance, fosters accountability, and strengthens the quality of public sector decision-making.

2. BACKGROUND

2.1. BASIC CONCEPTS

This section provides definitions of key terms used throughout the M&E Methodology Guide. These concepts are foundational for understanding different approaches, typologies, tools, and institutional arrangements related to monitoring and evaluation in the public sector.

The following table briefly outlines some of the key concepts that are mentioned throughout the guide.

TABLE 1: KEY CONCEPTS IN M&E

Concept	Definition
Monitoring	A continuous process of collecting, analyzing, and using data to track the implementation of interventions. It focuses primarily on inputs, activities, and outputs to assess whether implementation is proceeding as planned. Monitoring supports timely decision-making, course correction, and performance management.
Evaluation	A time-bound assessment that systematically examines the relevance, effectiveness, efficiency, coherence, sustainability, and impact of an intervention. Evaluation can be conducted at various stages (e.g., mid-term, final, ex-post) and is often used for accountability, learning, and strategic decision-making.
Inputs	The financial, human, and material resources mobilized for the implementation of an intervention. Inputs are the starting point in the results chain and are used to carry out planned activities.
Activities	The specific tasks or actions carried out using the inputs to deliver the intervention. Examples include conducting trainings, building infrastructure, developing software, or disbursing funds.
Outputs	The immediate products or services delivered as a result of activities. Outputs are usually tangible and directly attributable to the intervention (e.g., number of people trained, number of clinics built).
Outcomes	The short- to medium-term results expected from the use of outputs. Outcomes reflect changes in behavior, capacity, systems, or institutional performance, and can be influenced by factors beyond the intervention alone.
Impact	The long-term effects intended or unintended, positive or negative that the intervention contributes to at the population or systemic level. Impacts often relate to national development goals such as poverty reduction, employment, or environmental sustainability.
Results Chain	A logical model that describes how inputs are transformed into activities, which generate outputs, lead to outcomes, and contribute to impacts. Also known as a logic model or results framework.
Indicator	A quantitative or qualitative variable used to measure performance, change, or progress toward a result. Indicators must be specific, measurable, and time-bound. Each indicator should have a defined baseline, target, data source, and frequency of measurement.
Baseline	The starting value of an indicator before an intervention begins. It provides a point of reference for measuring change over time and evaluating impact.
Target	The specific value or level of achievement expected for an indicator at a defined point in time. Targets should be ambitious but realistic, and ideally set during planning stages.

Concept	Definition
Theory of Change (ToC)	A conceptual framework that explains how and why an intervention is expected to achieve its outcomes and impact. It identifies causal pathways, underlying assumptions, and external conditions. A ToC is typically visualized as a diagram and is supported by narrative explanation.
Logframe (Logical Framework)	A structured planning tool that presents the objectives, indicators, data sources, and assumptions of an intervention in a matrix format. While less flexible than a Theory of Change, it remains widely used in development programming.
SMART Indicators	A quality standard for indicator design. Indicators should be Specific, Measurable, Achievable, Relevant, and Time-bound. This ensures that indicators are meaningful and usable for monitoring and evaluation purposes.
Contribution Analysis	An evaluation approach that explores whether and how an intervention contributed to observed outcomes, particularly in complex settings where direct attribution is not possible. It tests a theory of change against multiple lines of evidence.
Attribution vs Contribution	Attribution refers to establishing a direct causal link between an intervention and an observed result (e.g., through counterfactuals). Contribution refers to demonstrating that an intervention plausibly influenced outcomes, alongside other factors. Most policy evaluations focus on contribution.
Rolling Evaluation Plan	A multi-year schedule of planned evaluations, updated annually, that helps prioritize evaluation efforts based on relevance, timing, and strategic value. Rolling plans support coordination, resource planning, and transparency.
Management Response	A formal reply by the institution responsible for an intervention, indicating whether and how it will address the recommendations of an evaluation. It is essential for institutional accountability and learning.
Regulatory Impact Assessment (RIA)	A structured method for assessing the actual (or anticipated) effects of a regulation. Ex-post RIA examines whether a regulation achieved its objectives, delivered public value, and remains relevant and proportionate.
Policy Impact Assessment (PIA)	An evaluation of whether a non-regulatory public policy has achieved its intended outcomes and produced meaningful change. PIA focuses on strategic initiatives, cross-sectoral programs, and public interventions without a direct legal component.

2.2. PRE-REQUISITES FOR A SUCCESSFUL M&E SYSTEM

Establishing a robust and effective Monitoring and Evaluation (M&E) system requires careful consideration of several foundational elements. These components are critical to ensuring the quality, reliability, and utility of monitoring and evaluation practices. The essential pre-requisite elements include:

Objectives and Principles: Clearly defined objectives and guiding principles are fundamental to the establishment of an effective M&E system. Objectives outline what the M&E system aims to achieve, while principles (e.g., accountability, transparency, and continuous learning) guide its implementation and operation.

Well-Defined Intervention Logic: A clear intervention logic is vital to delineating the causal relationships between inputs, activities, outputs, outcomes, and impacts. This logical structure provides a framework that underpins the entire M&E process, ensuring coherence in monitoring and evaluation efforts.

Baselines: Establishing baseline data is critical, as it provides the initial conditions against which future performance and outcomes are measured. Baselines should capture the situation at the onset of an intervention clearly and comprehensively.

Targets and Quality Standards: Setting explicit targets and quality standards enables effective tracking of performance. These standards and targets provide benchmarks for assessing whether an intervention is progressing as planned and meeting desired quality criteria.

Information Sources and Collection Methods: The identification of reliable data sources and clearly defined collection methods ensures consistency, accuracy, and credibility of monitoring and evaluation data. This involves selecting appropriate indicators, data gathering techniques, and data management systems.

Clear Definition of Tasks, Roles, and Responsibilities: A quality M&E system explicitly outlines the tasks, roles, and responsibilities of all stakeholders involved in data collection, analysis, reporting, and decision-making processes. Clearly delineated roles enhance accountability and efficiency.

Frequency of Measurements: Determining an appropriate measurement frequency is essential to ensure timely and useful data collection. The frequency should align with decision-making cycles, reporting needs, and the pace at which interventions are expected to produce measurable changes.

Timing of Evidence Gathering: Aligning the timing of data collection with key milestones in intervention implementation is crucial. Properly timed data collection enables informed decision-making, mid-course corrections, and timely reporting.

Reflection, Review, and Decision-making Processes: An effective M&E system includes structured processes for reflecting on collected evidence, reviewing progress, and making informed decisions. These processes promote continuous learning and adaptive management.

Communication Plan: A clearly articulated communication plan ensures the systematic dissemination of monitoring and evaluation findings to relevant stakeholders. Effective communication fosters transparency, enhances stakeholder engagement, and supports informed decision-making.

Adequate Resources for Monitoring and Reporting: Allocating sufficient resources – human, financial, and technical – is vital to support the ongoing operations of the M&E system. Adequate resources guarantee that monitoring activities and evaluation processes are consistently carried out, maintaining the system's quality and effectiveness.

Developing an effective and successful M&E system, does not happen in a vacuum. The monitoring and evaluation functions, are a key component of the policy cycle and should certainly be viewed within that lens. While an M&E framework does cover two key components, the planning component is also crucial, as it sets the stage for successful monitoring and evaluation. The primary role of the planning function in M&E is presented below, through the approach of Result Based Management.

2.2.1. POLICY PLANNING AS AN ENABLER FOR M&E: RESULTS BASED MANAGEMENT

Results-Based Management (RBM) provides the conceptual and practical foundation essential for systematically tracking and assessing the performance and impact of policies, programs, and projects. By clearly defining expected outcomes, RBM enhances accountability and fosters continuous learning and improvement. A central feature of RBM is the structured use of tools such as logical frameworks (logframes) and theories of change, which explicitly articulate the relationship between inputs, activities, outputs, outcomes, and impacts. The

figure below, outlines an example Result Chain, specific for Uzbekistan's digital economy training and SME grants.

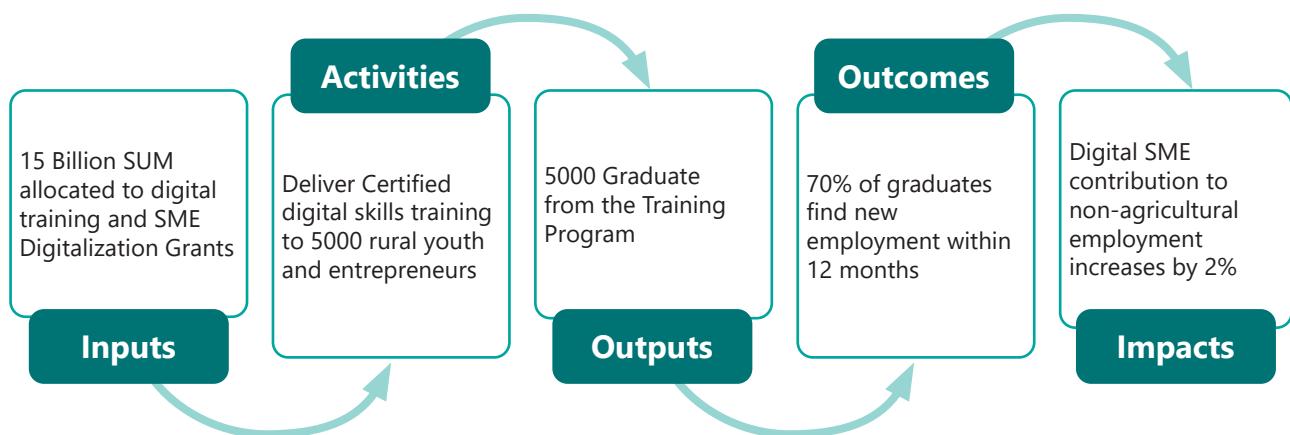


Figure 1: Example Result Chain

TOOLS FOR RBM: LOGICAL FRAMEWORKS AND THEORIES OF CHANGE

Logical frameworks and theories of change are pivotal tools for implementing RBM effectively. They:

- Clearly articulate the objectives and intended outcomes of interventions.
- Map transparent causal pathways from inputs through activities, outputs, outcomes, and ultimately to impacts.
- Facilitate the development of precise indicators essential for monitoring and evaluation.
- Offer structured approaches for monitoring progress, identifying implementation challenges, and supporting informed decision-making.

MULTI-LEVEL RESULTS CHAIN APPROACH

RBM employs a multi-level results chain to describe the progressive linkages between the resources invested and the ultimate impacts:

- **Inputs:** Resources invested in implementing a policy or program (e.g., human resources, financial resources, materials).
- **Activities:** Tasks undertaken utilizing these inputs to achieve policy or program objectives.
- **Outputs:** Direct, tangible products or services resulting from completed activities.
- **Outcomes:** Short- to medium-term effects derived from outputs, demonstrating meaningful progress toward strategic objectives.
- **Impacts:** Long-term, sustained changes in social, economic, environmental, or other conditions attributable directly or indirectly to the intervention.

Faoliyat va natijalar asosan monitoring uchun mavzulardir, chunki ular bezovsita nazorat va tezkor natijalarni o'z ichiga oladi. Natijalar va ta'sirlar, aksincha, kengroq, bilvosita tabiatli tufayli chuqurroq baholash tahlilini talab qiladi.

IMPORTANCE OF RESULTS FRAMEWORKS

A well-articulated results framework provides organizations with a unified vocabulary and a clear conceptual roadmap for managing portfolios. Such frameworks allow various stakeholders to have a common understanding and facilitate coherent planning and decision-making processes.

Moreover, results frameworks establish a crucial platform for the development of indicators, which serve as the cornerstone for effective monitoring and evaluation. By clearly defining expected outcomes and impacts, these frameworks enable precise measurement and reporting of progress, greatly enhancing the effectiveness and clarity of the entire M&E system.

2.2.2. DEVELOPMENT OF MONITORING PLAN

A Monitoring Plan provides the **strategic framework** for tracking the implementation and early results of policies, programs, and regulatory interventions. It defines **what will be monitored, how often, by whom, and using which sources of information**. A well-designed Monitoring Plan ensures that public institutions collect meaningful data to support management, accountability, and learning.

At the national level, the preparation of the overarching Monitoring Plan is the responsibility of the **Center of Government agency**¹ designated under the forthcoming M&E Decree. This agency will set minimum standards, define core indicators linked to national priorities, and ensure consistency across sectors. Its role is to ensure that monitoring efforts align with government-wide strategic goals and allow for performance aggregation across line ministries and programs.

However, the operationalization of monitoring meaning the day-to-day collection, management, and analysis of data rests primarily with **line Ministries, Departments, and Agencies (MDAs)**. Each MDA is responsible for designing sector- or program-specific monitoring frameworks, selecting appropriate indicators, setting baselines and targets, and ensuring that monitoring outputs are reliable and timely. MDAs must also ensure that monitoring systems are integrated into management processes, not treated merely as reporting requirements.

A strong Monitoring Plan should include:

- A clearly articulated results framework, connecting activities to outputs, outcomes, and impacts;
- A set of SMART indicators for each results level, linked to specific data sources;
- Baseline values and annual targets for tracking progress;
- Responsibilities for data collection, aggregation, and analysis;
- Frequency and methods of data collection (routine reporting, surveys, spot checks, MIS systems);
- Reporting schedules and formats to feed into sectoral and national performance reviews.

¹In this context, "Centre of Government Agency" refers to a government institution situated at the apex of the executive branch, undertaking the role of the Central Evaluation Coordinating Body.

Monitoring Plans should be considered living documents, updated as policies evolve, new indicators are required, or institutional capacity improves. The table below provides an overview of the Monitoring Plan development process:

TABLE 2: KEY STEPS IN DEVELOPING A MONITORING PLAN

Key Steps in Developing a Monitoring Plan	
1. Assign Institutional Responsibility	Clearly designate who is responsible for drafting and managing the monitoring plan. Each implementing institution should identify a monitoring focal point responsible for coordination and quality control.
2. Define What to Monitor	Identify the specific elements of the results framework that will be tracked. This includes: <ul style="list-style-type: none"> • Selecting indicators linked to inputs, activities, and outputs (as these fall within the control or direct influence of the implementers). • Prioritizing based on timing, feasibility, interest, and usefulness of the data. • Ensuring complementarity with existing systems or evaluations.
3. Determine Monitoring Type and Resource Needs	Choose the appropriate monitoring approach (e.g., routine data collection, rapid assessments, dashboard-based performance tracking) and assess human, financial, and technical resource needs. A minimum of 3–5% of the program budget should be dedicated to M&E.
4. Establish a Monitoring Schedule	Set a clear schedule that determines: <ul style="list-style-type: none"> • The timing and frequency of data collection (monthly, quarterly, annually) • The deadlines for validation, consolidation, and reporting • Milestones for reviewing the monitoring data and integrating it into decision-making.
5. Share and Validate the Monitoring Plan	Ensure the approved monitoring plan is shared with relevant stakeholders, including program managers, data providers, and oversight entities. Validation workshops may be useful to ensure relevance and buy-in.
6. Manage and Update the Plan	Monitoring plans are not static. They require ongoing coordination to: <ul style="list-style-type: none"> • Activate scheduled activities on time • Re-confirm availability of resources • Adjust indicators or formats if needed based on implementation realities.

MONITORING SYSTEM: KEY COMPONENTS AND ENABLERS

Apart from the existence and development of the plan, an effective Monitoring System depends on the following systemic enablers, and key components. These are outlined below:

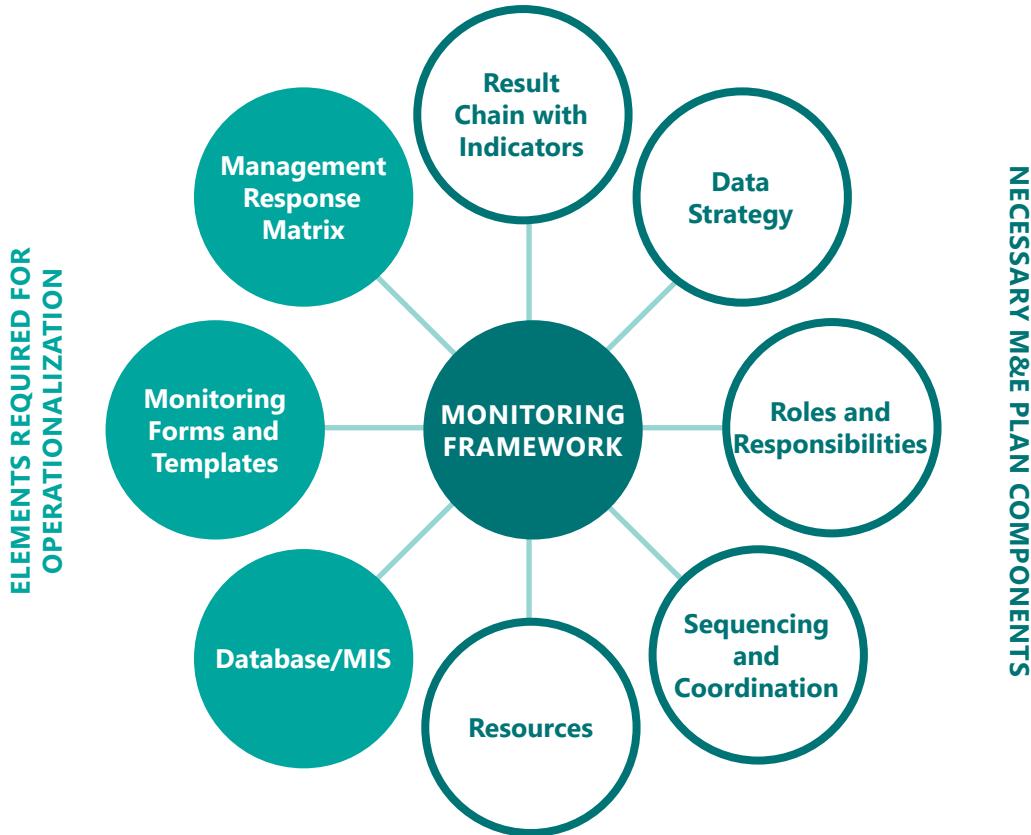


Figure 2: Key Components and Enablers for M&E Systems

- **Results chain with indicators:** Inputs, activities, outputs (and selected outcomes) with SMART indicators.
- **Data strategy:** Sources of data, tools to be used (forms, digital platforms), and storage systems.
- **Roles and responsibilities:** Who collects, who analyzes, who reports.
- **Sequencing and coordination:** Especially for multi-actor or joint monitoring efforts.
- **Resources:** Assigned budgets, staff, and technical tools.
- **Enabling Systems and Tools**

To operationalize a monitoring plan, the following elements are required:

- A structured database or Management Information System (MIS) to capture and process data
- Monitoring forms or templates that clearly define what information is to be collected and how
- A communication and dissemination strategy, outlining how monitoring information is used internally for learning and externally for accountability
- A management response matrix that records how M&E findings are being acted upon or followed up

Monitoring plans are most effective when directly linked to the Theory of Change and Results Framework. As such, they should be revisited periodically in light of evolving needs, policy changes, or implementation bottlenecks.

2.2.3. DEVELOPING OF EVALUATION PLAN

An Evaluation Plan defines the framework for systematically assessing the relevance, effectiveness, efficiency, coherence, impact, and sustainability of public interventions. It sets out when evaluations will be conducted, which interventions will be prioritized, what evaluation types will be used, and how findings will be utilized to support decision-making.

Similar to monitoring, the responsibility for preparing the **national-level Evaluation Plan** lies with the **Center of Government agency**. This Plan should cover a multi-year period (typically three to five years) and identify the priority areas for evaluation, based on criteria such as strategic importance, fiscal risk, public interest, innovation value, or policy relevance. The Center of Government ensures that evaluations are aligned with major policy cycles and budget planning, reducing duplication and addressing gaps across sectors.

The Evaluation Plan should outline:

- Priority interventions (policies, programs, regulations) selected for evaluation;
- Types of evaluation planned (ex-post RIA, ex-post PIA, program evaluations, thematic or sectoral evaluations);
- Expected timing and sequencing of evaluations (e.g., mid-term, final, or ex-post evaluations);
- Funding and human resource considerations (including provisions for independent evaluations);
- Quality assurance mechanisms, including peer reviews and validation processes;
- Expected use of findings for planning, reform, and strategic budgeting.

While the Center of Government coordinates the overall Evaluation Plan, **line MDAs** are responsible for the **design and implementation** of evaluations for the interventions they oversee. This includes developing Terms of Reference, managing the procurement or execution of evaluations (internal or external), facilitating access to data, and ensuring that recommendations are addressed through formal management responses.

Importantly, Evaluation Plans should not be static. They should allow for flexibility to commission evaluations on an ad-hoc basis in response to emerging priorities, unexpected developments, or requests from oversight bodies such as Parliament or the Supreme Audit Institution.

TABLE 3: KEY STEPS IN DEVELOPING AN EVALUATION PLAN

Key Steps in Developing an Evaluation Plan	
1. Strategic Prioritization	
Identify the priority areas or initiatives that merit evaluation within the planning period. This could be based on:	
<ul style="list-style-type: none"> • Strategic importance or alignment with national priorities • Visibility or public interest • Financial scale or risk level • History of prior evaluations (avoid duplication) 	
2. Timing and Sequencing	
Determine the most appropriate timing for evaluations based on:	
<ul style="list-style-type: none"> • The stage in the intervention cycle (ex-ante, mid-term, ex-post) • Alignment with planning, budgeting, or reporting cycles • Avoidance of conflicts with other institutional commitments 	
3. Evaluation Modalities and Types	
Select the most suitable evaluation type depending on the scope and purpose:	
<ul style="list-style-type: none"> • Formative vs. Summative • Project, Programmatic, Thematic, Sectoral, Country-wide, or Policy Evaluation • Impact Evaluation (where causal attribution is a priority) 	
4. Evaluation Criteria and Questions	
Define what the evaluation should assess using established criteria, especially the OECD-DAC criteria:	
<ul style="list-style-type: none"> • Relevance: Are the objectives still aligned with the needs of beneficiaries and national priorities? • Effectiveness: Were planned results achieved? • Efficiency: Were resources used in a cost-effective and timely way? • Impact: What higher-level effects (intended or unintended) were generated? • Sustainability: Will results likely continue beyond the intervention? 	
Frame the evaluation questions accordingly, and chose based on the purpose of the evaluation	
<ul style="list-style-type: none"> • Descriptive: What has happened? What has changed? • Causal: What led to the observed results? What was the intervention's contribution? • Normative: How do actual results compare to what was expected or desired? 	
5. Institutional Arrangements and Resources	
Assign clear roles for:	
<ul style="list-style-type: none"> • Evaluation management and coordination • Technical oversight and quality assurance • Internal vs. external evaluations 	
Include provisions for:	
<ul style="list-style-type: none"> • Budgeting (a portion of program funds should be allocated to evaluation) • Staff competencies and training 	
6. Integration with Monitoring Systems	
<ul style="list-style-type: none"> • Ensure baseline data and regular monitoring outputs feed into evaluation design. • Use monitoring trends to identify areas of concern or interest. 	
7. Follow-up and Use of Evaluation Findings	
Institutionalize a Management Response Mechanism to ensure evaluations lead to action. This includes:	
<ul style="list-style-type: none"> • Drafting and endorsing formal responses to evaluation recommendations • Assigning responsibilities and timelines for implementation • Tracking the uptake and use of evaluation results over time 	

EVALUATION ENABLING TOOLS AND PRACTICES

Apart from the plan, the following tools and practices ensure that an evaluation plan will be functioning and operational.

- **Rolling Evaluation Calendars:** These outline when evaluations will take place over a 3–5-year period and should be periodically updated. A rolling plan allows institutions to balance evaluation needs with available capacity and budget, while responding flexibly to emerging priorities.
- **Evaluation Management Templates:** Standardized formats such as Terms of Reference (ToRs), inception reports, evaluation matrices, and reporting templates help ensure consistency, quality, and comparability across evaluations. These tools are especially useful for institutions managing multiple evaluations across sectors.
- **Guidance Notes on Evaluation Standards:** Clearly defined quality assurance guidelines aligned with international norms (e.g., OECD-DAC, UNEG) provide a shared understanding of good evaluation practice. These notes may address evaluation ethics, data quality, triangulation methods, and reporting formats.
- **Stakeholder Involvement Frameworks:** Structured plans that define how different stakeholders – government units, beneficiaries, civil society, private sector – are engaged throughout the evaluation process. This improves credibility, relevance, and uptake of findings.
- **Evaluation Knowledge Management Platforms:** Digital repositories or databases to store, track, and disseminate evaluation reports, lessons learned, and management responses. These platforms support institutional memory and promote learning across the administration.
- **Management Response Mechanisms:** Tools that track how evaluation recommendations are being addressed, including accountability assignments and follow-up reporting cycles. These mechanisms ensure that evaluations lead to real-time decision-making and improvements.

When designed well, an evaluation plan transforms evaluation from a one-off activity into a **systematic learning and accountability tool** – embedded within public sector management processes and capable of informing strategic decision-making over time.

BOX 1: OECD DAC EVALUATION CRITERIA

Relevance: *The extent to which the intervention objectives and design respond to beneficiaries', global, country, and partner/institution needs, policies, and priorities, and continue to do so if circumstances change*

Efficiency: *The extent to which the intervention delivers, or is likely to deliver, results in an economic and timely way*

Effectiveness: *The extent to which the intervention achieved, or is expected to achieve, its objectives, and its results, including any differential results across groups*

Impact: *The extent to which the intervention has generated or is expected to generate a significant positive or negative, intended or unintended, higher-level effects*

Sustainability: *The extent to which the net benefits of the intervention continue, or are likely to continue*

3. EVALUATION TYPOLOGY

Evaluation, at its core, is a **structured** yet **flexible** inquiry process one that allows public administrations to make sense of whether and how their interventions work. Unlike compliance-based monitoring, which focuses on tracking implementation against pre-set indicators, **evaluation embraces complexity**. It asks **deeper questions** about results, mechanisms of change, contextual dynamics, and value for money. As such, evaluation is not simply a bureaucratic obligation; it is an inherently creative and adaptive function, capable of responding to a broad range of learning, accountability, and strategic needs.

The strength of evaluation lies in its **capacity to accommodate diversity in methods**, in **questions**, and in the **types of interventions being assessed**. A single evaluation might examine a regulation's unintended effects, a national strategy's relevance to evolving development goals, or the long-term impact of a digital inclusion program. This diversity is captured through multiple **typologies**, which help practitioners structure evaluations while retaining flexibility.

These typologies reflect different entry points into the evaluation process. **Use orientation** (e.g., formative, summative, anticipatory) clarifies what the evaluation is meant to support whether learning, decision-making, accountability, or strategic foresight. **Timing** addresses when the evaluation takes place in the intervention lifecycle: before (ex-ante), during (mid-term), or after (final or ex-post). **Subject** typologies classify evaluations by what is being assessed – regulations (via RIA), policies (via PIA), or programs and projects (via RBM). **Scope or scale** determines how broad or granular the evaluation should be ranging from single interventions to sectoral reviews or system-level meta-evaluations.

While the tools and approaches vary, **the core evaluation criteria remain consistent**. Across nearly all evaluation types, practitioners are expected to assess relevance, efficiency, effectiveness, impact, and sustainability most commonly through the OECD-DAC framework. However, the **questions guiding the evaluation must be tailored** to its purpose. For instance, a regulatory impact evaluation (RIA) conducted ex-ante might ask: *Are there viable non-regulatory alternatives?* or *What is the likely compliance burden?* In contrast, a policy impact assessment (PIA) may focus on *policy coherence, trade-offs, or effects across population groups*. RBM-linked evaluations are typically more operational in nature, assessing whether activities and outputs translate into tangible outcomes and, ultimately, sustainable impact.

Each of these methodologies RIA, PIA, and RBM offers a structured yet distinct approach:

- **Regulatory Impact Assessments (RIA)** are typically anticipatory and ex-ante in nature, focused on proposed regulatory interventions. Their aim is to forecast potential costs, benefits, compliance risks, and behavioral impacts before a regulation is enacted. RIA supports decision-makers in considering alternatives, refining regulatory proposals, and ensuring legal proportionality.
- **Policy Impact Assessments (PIA)** operate at the level of government strategies and broad policy frameworks. They combine ex-ante design validation, interim monitoring, and ex-post evaluation. PIA is inherently interdisciplinary, integrating political economy analysis, stakeholder mapping, and implementation diagnostics. It's particularly valuable in complex, cross-sectoral policies – such as national digital transformation plans or climate strategies.

■ **Results-Based Management (RBM)** focuses on measurable change across the intervention logic – from inputs and activities to outputs, outcomes, and impacts. It provides the foundation for program and project-level M&E by establishing results frameworks, logical models, and clear performance indicators. RBM evaluations are typically mid-term (formative) or final (summative), and they serve both learning and accountability functions.

Despite this typological richness, it is **neither necessary nor desirable to codify every evaluation sub-type into legislation or institutional frameworks**. Instead, a good M&E system should provide evaluators with structured **guidance**, criteria, typologies, and process standards, while leaving room for professional discretion. **Evaluators must be empowered to design inquiries** around what really matters: the **learning and accountability needs of the institutions**, the complexity of the intervention, and the strategic relevance of the questions at hand.

For example, if a ministry is implementing a high-stakes employment reform in response to economic shocks, the most valuable evaluation may not be a traditional summative review. Instead, a series of lighter, formative evaluations paired with periodic RBM dashboards and a final impact assessment could yield far more actionable knowledge. Similarly, while a national law on digital access may warrant a formal RIA prior to approval, a follow-up PIA might better capture its social equity effects once implemented.

The following matrix presents a **structured typology of evaluation types used in public sector M&E**. It provides a comparative view of how evaluations can differ by purpose, timing, scope, and object. This typology is designed to **help practitioners select the most appropriate evaluation type** based on the nature of the intervention, the timing of the assessment, and the institutional context. Each type is paired with typical questions and tools used, allowing for better alignment between evaluative purpose and method

4-JADVAL: BAHOLASH TIPOLOGIYALARI

Evaluation Type	Purpose/Use Orientation	Timing	Scale	Subject/Approach	Typical Tools and Questions
Anticipatory (Ex-ante)	Inform design, assess feasibility	Pre-implementation	Project, Policy	RIA, ex-ante PIA	CBA, logic models, stakeholder mapping; "What might happen?"
Formative	Adaptation, real-time learning	During implementation	Project, Program	RBM, PIA	TOC validation, feedback loops; "What's working? What needs adjustment?"
Summative (Final)	Accountability, overall results	End of implementation	Any scale	RBM, PIA	Endline surveys, KPI analysis; "Did we achieve our outcomes? Was it worth the cost?"
Ex-post	Sustainability, long-term effects	Years after completion	Project, Policy, Sector	Impact Evaluation	Tracer studies, contribution analysis; "What has endured? What were the lasting effects?"
Compliance/Performance	Routine accountability, standards	Continuous	Project, Ministry-wide	RBM	Dashboards, scorecards; "Are we meeting obligations? Are KPIs on track?"
Policy Evaluation	Policy effectiveness & coherence	Mixed (often ex-post)	National/sectoral	PIA	Policy coherence analysis, contribution analysis; "Is the policy solving the issue?"

Evaluation Type	Purpose/Use Orientation	Timing	Scale	Subject/Approach	Typical Tools and Questions
Program Evaluation	Assess program logic and impact	Mid/final	Thematic/clustered	RBM, RIA	Cluster evaluation, outcome mapping; "Are the grouped interventions effective together?"
Thematic Evaluation	Cross-cutting issues, learning	Mixed	Multi-sector, national	Any	Portfolio review, synthesis tools; "What's our performance on gender/youth/etc.?"
Sectoral Evaluation	Strategic planning, system performance	Mid/ex-post	Sector-wide	Mix of RIA, PIA, RBM	Sector scans, strategic reviews; "How is the sector performing overall?"
Impact Evaluation	Causality and attribution	Final or ex-post	Usually project-level	Any	RCTs, counterfactuals, contribution analysis; "What change can be attributed to this?"
Meta-Evaluation	System learning, evaluation quality	Any time	National/system-wide	Evaluation System	Synthesis grids, quality reviews; "Are our evaluations reliable and useful?"
Economic Evaluation	Value-for-money, ROI analysis	Usually ex-ante/final	Project/Policy	RIA, RBM	CBA, SROI, ROI calculators; "Are the costs justified by the benefits?"

4. EVALUATION METHODS

Given the overall evaluation typology from the above table, this section presents some of the key evaluation types alongside their key features, core steps, and common tools used.

TABLE 5: EX-ANTE RIA

Ex-Ante Regulatory Impact Assessment (RIA)	
Used before a regulation is enacted, ex-ante RIA is a predictive analytical tool designed to inform whether regulatory intervention is justified and, if so, how to design it most effectively and efficiently.	
Key Features:	
<ul style="list-style-type: none"> Applies to draft regulations and legal acts with societal or economic impacts. Supports transparent decision-making by analyzing regulatory alternatives. Helps avoid unnecessary administrative burden or policy failure. 	
Core Steps:	
<ul style="list-style-type: none"> Problem Identification: Define the policy issue and its root causes. Objective Setting: Outline the regulation's purpose and expected outcomes. Option Identification: Identify regulatory and non-regulatory alternatives. Impact Assessment: Estimate costs, benefits, risks, and distributional effects. Stakeholder Engagement: Validate assumptions and capture feedback. Preferred Option Selection: Justify the selected approach. RIA Statement: Prepare a clear summary for decision-makers. 	
Common Tools	
Cost-Benefit Analysis, Risk Matrices, Administrative Burden Calculators, Multi-Criteria Decision Analysis.	

TABLE 6: EX-POST RIA

Ex-Post Regulatory Impact Assessment (RIA)	
Conducted after the regulation has been in force, ex-post RIA evaluates whether the regulation achieved its intended effects and whether it remains relevant, efficient, and proportionate.	
Key Features:	
<ul style="list-style-type: none"> Focuses on actual implementation outcomes versus expected ones. Informs amendment, simplification, or repeal of regulations. Supports regulatory stock management and burden reduction. 	
Core Steps:	
<ul style="list-style-type: none"> Establish Evaluation Questions: Draw from original RIA and current concerns. Define Criteria: Relevance, effectiveness, efficiency, proportionality, coherence. Data Collection: Use administrative data, stakeholder input, and compliance records. Impact Assessment: Estimate costs, benefits, risks, and distributional effects. Stakeholder Engagement: Validate assumptions and capture feedback. Gap and Burden Analysis: Identify unintended consequences or overregulation. Recommendations: Propose improvements or de-regulation options 	
Common Tool	
<ul style="list-style-type: none"> Compliance Monitoring, Stakeholder Surveys, KPI Tracking, Post-Implementation Reviews. 	

TABLE 7: EX-ANTE PIA

Ex-Ante Policy Impact Assessment (PIA)	
Ex-ante PIA is used prior to the adoption of policies or strategies to explore potential effects, alignment with strategic goals, and delivery feasibility.	
Key Features:	
<ul style="list-style-type: none"> • Applied to policy concepts or draft strategies, especially large-scale or cross-sectoral. • Focuses on design optimization and problem-solution fit. • Helps avoid underperforming or misaligned policy interventions 	
Core Steps:	
<ul style="list-style-type: none"> • Strategic Framing: Define the policy problem and intended outcomes. • Theory of Change: Map expected causal pathways. • Stakeholder Mapping: Identify affected groups and institutions. • Scenario Analysis: Explore policy alternatives and their implications. • Feasibility Testing: Assess institutional, financial, and political viability. • Option Comparison and Justification 	
Common Tools	
Compliance Monitoring, Stakeholder Surveys, KPI Tracking, Post-Implementation Reviews.	

TABLE 8: EX-POST PIA

Ex-Post Policy Impact Assessment (PIA)	
Ex-post PIA assesses a policy after implementation to determine whether it delivered the expected outcomes and what lessons can inform future policies.	
Key Features:	
<ul style="list-style-type: none"> • Supports policy learning and mid-cycle corrections. • Essential for high-impact, high-visibility policies. • Informs renewal, expansion, or discontinuation 	
Core Steps:	
<ul style="list-style-type: none"> • Define Evaluation Scope and Questions: Based on policy goals and logic. • Collect Evidence: Using both monitoring data and new evaluations. • Analyze Results: Look at effectiveness, equity, sustainability, and coherence. • Stakeholder Consultations: Gather perspectives from beneficiaries and implementers. • Distill Lessons: What worked, for whom, under what conditions? • Inform Future Policy Cycles. 	
Common Tools	
Contribution Analysis, Outcome Harvesting, Policy Coherence Reviews, Before-After Comparisons, Disaggregated Impact Reviews.	

TABLE 9: THEMATIC EVALUATION

Thematic Evaluation (Primarily Ex-Post or Mid-Term)	
To assess a specific theme or cross-cutting issue that spans across multiple interventions, sectors, or geographic areas (e.g., gender equality, digitalization, climate resilience).	
Key Features:	
<ul style="list-style-type: none"> • Can include programs, policies, and regulations across institutions. • Focuses on coherence, synergies, gaps, and results within a theme. • Often commissioned centrally or by donors to support strategic planning. 	
Core Steps:	
<ul style="list-style-type: none"> • Define the theme and its relevance to national priorities. • Select a representative sample of interventions (programs, projects, regulations). • Develop a theory of change or results map for the thematic area. • Analyze evidence across cases: outputs, outcomes, sustainability, coherence. • Formulate system-wide conclusions and recommendations. 	
Common Tools	
Outcome Mapping, Comparative Case Studies, Meta-Synthesis, Contribution Analysis.	

TABLE 10: SECTORAL EVALUATION

Sectoral or Strategic Evaluation (Usually Ex-Post or Periodic)	
To assess the performance of an entire sector (e.g., health, education, agriculture) or strategic pillar (e.g., public administration reform), typically to inform strategy development, restructuring, or investment planning.	
Key Features:	
<ul style="list-style-type: none"> • Evaluates policies, programs, and institutions in aggregate. • Often aims to align sector interventions with national or international targets (e.g., SDGs). • Useful for ministries and planning units. 	
Core Steps:	
<ul style="list-style-type: none"> • Map all interventions and institutional actors in the sector. • Define strategic and performance indicators. • Assess alignment, coherence, and performance against sector goals. • Identify systemic bottlenecks, policy gaps, and best practices. • Recommend sector reform options or program reprioritization. 	
Common Tools	
Sector Performance Reviews, SWOT, Policy Coherence Analysis, Budget Efficiency Studies.	

TABLE 11: CLUSTER EVALUATION

Cluster Evaluation
To assess a group of related interventions (e.g., all digital transformation programs; all MSME support projects), typically to understand collective contribution to shared outcomes.
Key Features:
<ul style="list-style-type: none"> • Less about geographic or thematic breadth, more about programmatic clustering. • Enables cross-learning between similar projects or programs. • Often supports decisions on scaling up or consolidating efforts.
Core Steps:
<ul style="list-style-type: none"> • Define the cluster logic (thematic, donor portfolio, target group, etc.). • Review each intervention individually. • Identify overlaps, synergies, and points of divergence. • Analyze aggregated results and cost-effectiveness. • Draw cross-cutting conclusions and recommend strategic actions.
Common Tools
Portfolio Review, Logical Model Comparison, Synthesis Grids.

TABLE 12: PROJECT EVALUATION

Project Evaluation
To evaluate a single project and its performance in delivering defined outputs and outcomes.
Key Features:
<ul style="list-style-type: none"> • Most common evaluation form. • Focuses on relevance, implementation, efficiency, and immediate outcomes. • Used heavily in donor-funded or technical assistance settings.
Core Steps:
<ul style="list-style-type: none"> • Reconstruct the project logic and results framework. • Assess implementation fidelity and operational performance. • Evaluate achievement of short- and medium-term results. • Review management and resource use. • Provide lessons and improvement areas.
Common Tools
Logframes, Endline Surveys, Output-Outcome Tracking, ToC Reviews.

TABLE 13: PROGRAM EVALUATION

Program Evaluation
To assess a set of linked interventions (projects or actions) that together pursue a common objective (e.g., a national employment program with several delivery mechanisms).
Key Features:
<ul style="list-style-type: none"> • Aggregates across modalities but under a unified strategy. • May assess design integration, management coordination, and value for money. • Important for donor harmonization and national budget programs.
Core Steps:
<ul style="list-style-type: none"> • Clarify the program's scope and internal logic. • Review component interventions. • Assess horizontal integration (institutional coherence). • Evaluate program-level effectiveness and outcomes. • Identify duplications, overlaps, and missed synergies.
Common Tools
Results Matrices, KPI Reviews, Budget and Output Aggregation, Stakeholder Mapping.

TABLE 14: POLICY EVALUATION

Policy Evaluation
To evaluate the performance and impact of an implemented public policy, particularly its outcomes, relevance, and alignment with broader objectives.
Key Features:
<ul style="list-style-type: none"> • Applies to national or sector policies (not legislative regulations). • Focus on strategy coherence, equity, and institutional performance. • Typically large-scale, qualitative and quantitative mix.
Core Steps:
<ul style="list-style-type: none"> • Revisit original objectives and assumptions. • Map implementation and policy actors. • Analyze outputs, outcomes, and unintended effects. • Assess fit-for-purpose, alignment, and trade-offs. • Draw system-level and operational conclusions.
Common Tools
Policy Coherence Reviews, Contribution Analysis, Strategic Foresight for Relevance Testing.

TABLE 15: IMPACT EVALUATION

Impact Evaluation
To rigorously assess the long-term effects of an intervention, particularly to establish causality – i.e., whether observed outcomes can be attributed to the intervention itself.
Key Features:
<ul style="list-style-type: none"> • Focuses on attribution vs. contribution. • Requires high methodological rigor and baseline data. • Often involves control groups or sophisticated counterfactual models.
Core Steps:
<ul style="list-style-type: none"> • Develop a theory of change. • Define target outcomes and counterfactual. • Select evaluation design (RCT, quasi-experimental, etc.). • Collect baseline and follow-up data. • Analyze changes attributable to the intervention.
Common Tools
RCTs, Difference-in-Differences, Propensity Score Matching, Tracer Studies.

TABLE 16: META EVALUATION

Meta Evaluation (or system-level evaluation)
To rigorously assess the long-term effects of an intervention, particularly to establish causality – i.e., whether observed outcomes can be attributed to the intervention itself.
Key Features:
<ul style="list-style-type: none"> • A second-order evaluation. • Used to improve evaluation practice, coverage, and institutional learning. • Can synthesize evaluation findings across an organization or sector
Core Steps:
<ul style="list-style-type: none"> • Compile relevant evaluations from defined time frame. • Apply quality criteria (methodological soundness, usefulness, etc.). • Analyze patterns in findings, gaps, strengths. • Recommend institutional reforms and learning processes.
Common Tools
Evaluation Quality Rubrics, Synthesis Templates, Evaluation Scorecards.

4.1. TOOLS FOR EVALUATION

Evaluation tools are the specific instruments, techniques, and templates used to support the application of evaluation methods. While methods provide the overall design or strategy for the evaluation, tools help implement that strategy for example, by structuring data collection (surveys, interview guides), organizing information (evaluation matrices, contribution maps), or visualizing findings (dashboards, scorecards). Tools vary depending on the evaluation type, sector, and scope, and many are shared across monitoring and evaluation processes. The following table provides an overview of key tools used in policy, program, and regulatory evaluations, organized by their primary purpose and typical point of use in the evaluation cycle.

TABLE 17: COST-BENEFIT ANALYSIS

Cost-Benefit Analysis (CBA)	
Cost-Benefit Analysis is a systematic approach to evaluating the economic pros and cons of different options, aiding decision-makers in determining the most beneficial course of action	
Key Components:	
<ul style="list-style-type: none"> Identification of Costs and Benefits: List all potential costs and benefits associated with the project or decision. Monetary Valuation: Assign a monetary value to each identified cost and benefit. Time Horizon: Define the period over which the costs and benefits will be assessed. Discounting: Adjust future costs and benefits to present value using a discount rate. Net Present Value (NPV): Calculate the difference between the present value of benefits and costs. Sensitivity Analysis: Assess how results change with variations in key assumptions or variables. 	
Process:	
<ul style="list-style-type: none"> Define the Scope: Clearly outline the project or decision context Identify Costs and Benefits: Catalogue all relevant costs and benefits, both tangible and intangible Quantify and Monetize: Measure and assign monetary values to each cost and benefit Discount Future Values: Convert future amounts to present value using an appropriate discount rate. Calculate NPV: Subtract total discounted costs from total discounted benefits. Perform Sensitivity Analysis: Test the robustness of the results under different scenarios. Make a Recommendation: Based on the NPV and analysis, advise on the feasibility of the project or decision 	
Benefits:	
<ul style="list-style-type: none"> Facilitates objective comparison between alternatives. Incorporates a comprehensive view of economic impacts. Supports transparent and informed decision-making. 	
Challenges:	
<ul style="list-style-type: none"> Difficulty in quantifying intangible benefits and costs. Potential biases in estimating future values. Selecting an appropriate discount rate can be complex. 	

TABLE 18: THEORY OF CHANGE

Theory of Change (ToC)
Theory of Change is a comprehensive methodology used to describe how and why a desired change is expected to happen in a particular context.
Key Components:
<ul style="list-style-type: none"> • Long-Term Goals: Define the ultimate objectives or impacts desired. • Outcomes Pathway: Map the intermediate and early outcomes that lead to the long-term goals. • Interventions: Identify the activities or strategies implemented to achieve the outcomes. • Assumptions: Articulate the underlying beliefs about how change will occur. • Indicators: Establish metrics to measure progress toward outcomes.
Process:
<ul style="list-style-type: none"> • Engage Stakeholders: Involve all relevant parties to ensure diverse perspectives. • Define Long-Term Goals: Clearly state the desired change or impact. • Map Outcomes Pathway: Illustrate the sequence of outcomes leading to the goal. • Identify Interventions: Determine the actions required to achieve each outcome. • Surface Assumptions: Make explicit the assumptions underpinning the change process. • Develop Indicators: Create measurable signs of progress for each outcome. • Document and Review: Compile the ToC and regularly revisit it to ensure relevance.
Benefits
<ul style="list-style-type: none"> • Provides a clear roadmap linking activities to desired outcomes. • Enhances strategic planning and evaluation. • Facilitates stakeholder alignment and communication.
Challenges
<ul style="list-style-type: none"> • Can be time-consuming to develop comprehensively. • Risk of oversimplifying complex change processes. • Requires continuous updating as contexts evolve

TABLE 19: ROI ANALYSIS

Return on Investment (ROI) Analysis
ROI analysis evaluates the profitability of an investment by comparing the net benefits to the costs, expressed as a percentage.
Key Components:
<ul style="list-style-type: none"> • Investment Costs: Total expenses incurred for the investment. • Net Benefits: Total gains minus total costs. • Time Frame: Period over which the investment's returns are evaluated.
Process:
<ul style="list-style-type: none"> • Identify Costs and Benefits: List all relevant expenditures and gains. • Quantify Financial Values: Assign monetary values to costs and benefits. • Calculate ROI: Use the formula: $ROI = (\text{Net Benefits} / \text{Investment Costs}) \times 100\%$. • Interpret Results: Assess the profitability indicated by the ROI percentage
Benefits
<ul style="list-style-type: none"> • Simplifies profitability assessment. • Facilitates comparison between investment options
Challenges
<ul style="list-style-type: none"> • May overlook non-financial benefits • Does not account for the time value of money

TABLE 20: SECTOR SCAN

Sector Scan	
<p>A sector scan is a structured tool used in public policy and planning to map the current landscape of a policy sector (e.g., health, education, social protection, digital economy). It helps identify the main actors, interventions, financing flows, regulatory frameworks, and performance trends. The tool is typically used during strategic planning, evaluation, or reform design phases to support evidence-based policy choices.</p>	
Key Components: <ul style="list-style-type: none"> Sector Overview: Define the policy sector's mandate, scope, and institutional landscape. Policy and Regulatory Environment: Identify core laws, strategies, and governance mechanisms guiding the sector. Programmatic Inventory: Map current and recent interventions, projects, and funding envelopes. Results Landscape: Identify key achievements, outcome trends, gaps, and indicators tracked. Coordination Architecture: Analyze inter-ministerial, donor, and stakeholder coordination mechanisms. Capacity and Resource Assessment: Review human, institutional, and financial capacities of implementing bodies. 	
Process: <ul style="list-style-type: none"> Define Objectives and Scope: Clarify what the scan is meant to inform (e.g., sector strategy, reform, program evaluation). Gather Data: Use national plans, sector reviews, budget documents, and stakeholder interviews. Synthesize Findings: Analyze strengths, weaknesses, opportunities, and threats across the sector. Identify Strategic Gaps: Pinpoint performance bottlenecks, underfunded areas, or coordination failures. Formulate Insights: Generate forward-looking recommendations for policy improvement or strategic investment. 	
Benefits <ul style="list-style-type: none"> Enables a strategic understanding of sector dynamics for planners and evaluators. Supports evidence-based prioritization of reforms and investments. Enhances coherence across multiple interventions or donors in the same sector. 	
Challenges <ul style="list-style-type: none"> Requires integration of disparate data sources (finance, performance, institutions). May uncover politically sensitive misalignments or institutional overlaps. Effectiveness depends on the quality and timeliness of underlying data. 	

TABLE 21: PORTFOLIO REVIEW

Portfolio Review
A portfolio review systematically evaluates a set of related interventions (projects, policies, or programs) to identify overall trends, effectiveness, coherence, and alignment with strategic objectives.
Key Components:
<ul style="list-style-type: none"> Scope Definition: Clearly define interventions included in the portfolio. Performance Criteria: Establish standards for assessing portfolio performance (e.g., relevance, effectiveness, coherence). Comparative Analysis: Assess how different interventions perform relative to each other. Strategic Alignment: Evaluate alignment with broader organizational or governmental goals.
Process:
<ul style="list-style-type: none"> Define the portfolio's scope, objectives, and evaluation criteria. Collect and aggregate data from individual projects or programs. Conduct comparative performance analysis. Analyze strategic alignment and coherence. Report findings and recommend strategic adjustments.
Benefits
<ul style="list-style-type: none"> Identifies synergies and redundancies among interventions. Enhances resource allocation efficiency. Improves strategic decision-making and coordination.
Challenges
<ul style="list-style-type: none"> Managing large volumes of diverse data. Ensuring consistent evaluative standards across interventions. Potential complexity in synthesizing divergent results.

TABLE 22: INSTITUTIONAL ASSESSMENT

Institutional Assessment
An institutional assessment is a structured evaluation of a public body's internal capacities, performance, and alignment with its mandate and strategic objectives. It helps identify strengths, gaps, and areas for organizational improvement in relation to service delivery, policy implementation, and cross-institutional coordination.
Key Components:
<ul style="list-style-type: none"> Mandate and Strategic Clarity: Assess whether the institution's functions, authority, and strategic priorities are well-defined and up to date. Governance and Leadership: Examine internal governance structures, leadership effectiveness, and decision-making processes. Performance Management: Review how the institution monitors outputs and outcomes, and how it responds to performance gaps. Human and Technical Capacity: Evaluate staffing adequacy, skills, and the availability of technical tools and systems. Financial and Resource Management: Assess efficiency, transparency, and adequacy of budgeting and financial controls. Coordination and Accountability: Analyze institutional interactions (e.g., with line ministries, oversight bodies) and mechanisms for public accountability.

Process:
<ul style="list-style-type: none"> • Define Scope and Objectives: Determine whether the assessment focuses on strategic alignment, operational performance, reform readiness, or institutional resilience. • Data Collection: Use internal documents, organograms, performance reports, staff surveys, and key informant interviews. • Capability and Gap Analysis: Benchmark against standards, peer institutions, or international frameworks. • Synthesis of Findings: Highlight capacity bottlenecks, systemic risks, and areas of strength.
Benefits
<ul style="list-style-type: none"> • Supports reform planning, capacity development, or restructuring processes. • Enables targeted follow-up through organizational improvement plans. • Promotes institutional learning and greater internal coherence.
Challenges
<ul style="list-style-type: none"> • Requires access to sensitive performance and HR data. • Outcomes can be politically sensitive if linked to structural change. • Depends on the institution's willingness to engage honestly and act on findings.

TABLE 23: ENDLINE SURVEY

Endline Surveys
Endline surveys measure the final status of targeted indicators after an intervention has been implemented, enabling comparison with baseline data to assess effectiveness.
Key Components:
<ul style="list-style-type: none"> • Indicator Selection: Clearly defined outcome indicators based on intervention objectives. • Sampling Strategy: Representative sampling to ensure data reliability. • Comparison with Baseline: Systematic assessment of changes from initial conditions. • Statistical Analysis: Rigorous analysis to attribute changes to the intervention.
Process:
<ul style="list-style-type: none"> • Define clear objectives and indicators to measure. • Develop a survey instrument aligned with baseline measures. • Conduct data collection following standardized protocols. • Analyze data, comparing endline with baseline results. • Report findings, emphasizing changes attributed to interventions.
Benefits
<ul style="list-style-type: none"> • Provides evidence of intervention effectiveness and impact. • Supports accountability and informed decision-making. • Enables quantifiable measurement of change over time.
Challenges
<ul style="list-style-type: none"> • Data quality consistency between baseline and endline. • Attribution difficulties if external factors influenced outcomes. • Potential logistical complexities in data collection.

4.2. PROCESS FOR SELECTING EVALUATION TYPE

Public sector evaluations serve many functions informing decisions, assessing results, supporting accountability, and enabling system-wide learning. However, evaluations are not all the same. Choosing the **right evaluation type** requires a structured and informed approach, one that matches the characteristics of the intervention, the stage in its lifecycle, and the purpose behind the evaluation.

The process described below is designed to help practitioners navigate that decision logically, while retaining space for professional judgment and contextual nuance. It draws on the full typology of evaluation types developed under Uzbekistan's M&E reform framework.

4.2.1. STEP 1: IDENTIFY THE EVALUATION TRIGGER

The first step is deceptively simple – but essential: **Why is this evaluation happening?** Evaluations can be triggered by a range of formal or informal drivers, each carrying different implications for scope, timing, and independence.

In some cases, the evaluation is **legally or procedurally mandated**. For example, a new regulation may require an **ex-ante Regulatory Impact Assessment (RIA)** before it can be enacted, or a law may require its effects to be reviewed every five years – triggering an **ex-post RIA**. In other cases, the evaluation is **built into the intervention design**. Many large-scale programs or sectoral strategies include a **mid-term review** or **final evaluation** in their results frameworks.

But not all evaluations are anticipated. Some arise reactively – triggered by poor performance, policy drift, leadership changes, or political crises. These **strategic or management-driven evaluations** are often commissioned to guide reform or explain underachievement. Others come from the **donor or oversight community**, such as when multiple programs are pooled into a **cluster evaluation**, or when a central agency commissions a **meta-evaluation** of the entire evaluation system.

Tip: Understanding the trigger helps anticipate how flexible the evaluation can be. A donor-mandated summative evaluation will require different stakeholder management than a ministry-led learning review.

4.2.2. STEP 2: DEFINE WHAT IS BEING EVALUATED

Having clarified the trigger, the next step is to pin down the **object of evaluation**. Without this, evaluations risk becoming abstract or misdirected.

Practitioners must ask: *What, specifically, is the focus of inquiry?* Is it a regulation (a legally binding act)? A policy or strategy (a statement of intent and direction)? A program (a set of services or activities with budgeted resources)? A project (a time-bound intervention with specific outputs)? Or is the evaluation examining something broader – such as a **portfolio of interventions**, a **whole policy sector** (like education or health), or a **thematic area** (like gender or digital transformation)?

Each of these objects aligns with specific evaluation types. For example:

- A **regulation** may require a **RIA** (ex-ante) or **ex-post review** focused on compliance or burden.

- A **policy** could be the subject of a **Policy Impact Assessment (PIA)**, or an **ex-post policy evaluation** to assess outcomes and coherence.
- A **program** might be evaluated through **formative or summative program evaluations**, depending on timing.
- A **sector** might undergo a **sectoral evaluation** to assess strategic fit, institutional coherence, and results across agencies.
- A **thematic area** – like climate or youth – is best served through a **thematic evaluation**, often drawing from multiple programs and data sources.

Keep in Mind: Evaluations often begin with an unclear or overly broad object (for example the youth entrepreneurship sphere) This must be refined into a concrete unit (e.g., youth employment regulation, youth entrepreneurship program) to be meaningful.

4.2.3. STEP 3: DETERMINE THE TIMING IN THE LIFECYCLE

Once the object is defined, the next critical step is to ask: *Where are we in the lifecycle of this intervention?*

Timing dictates the types of questions that can be answered, the data that is available, and the evaluation's role in decision-making.

- If the evaluation is being conducted **before implementation begins**, it is an **ex-ante evaluation**. This includes **ex-ante RIA** (for regulations) and **PIA** (for policies). These are forward-looking and support decision-makers by forecasting impacts, costs, risks, and benefits.
- If the intervention is **mid-way through implementation**, a **mid-term review or formative evaluation** may be appropriate. These focus on operational learning, course correction, and stakeholder feedback. They are common in large public programs or multi-year national plans.
- If the intervention has **recently or long ago concluded**, an **ex-post evaluation** is needed. This could take the form of an **impact evaluation**, **policy evaluation**, or **sectoral review**, depending on the object and purpose.

Example: A national digital inclusion policy may require an ex-ante PIA to assess feasibility and a later ex-post evaluation to examine equity impacts and digital literacy outcomes.

Common Pitfall: Final evaluations are often commissioned before sufficient results have matured. This leads to superficial assessments and missed learning opportunities. If timing is uncertain, consider conducting an **evaluability assessment** first.

4.2.4. STEP 4: CLARIFY THE PURPOSE OR INTENDED USE

Even with the object and timing determined, practitioners must clarify the **intended use** of the evaluation. *What is this evaluation meant to do?* Is it to inform design? Improve delivery? Demonstrate results? Guide strategic planning?

Common purposes include:

- **Design Support:** Common in **ex-ante RIAs** and **PIAs**, helping prevent policy failure before it starts.
- **Learning and Adaptation:** Often mid-term or formative, guiding ongoing improvements in **programs** or **policies**.
- **Accountability and Reporting:** Final evaluations and **impact assessments** often serve this role, especially when donors or parliaments are involved.
- **Strategic Alignment:** **Sectoral, thematic**, or **cluster evaluations** are commissioned to assess coherence and prioritization across multiple interventions.
- **System Learning:** **Meta-evaluations** review how evaluation systems function and where institutional improvements are needed.

Tip: *Trying to serve too many purposes in one evaluation (e.g., learning + accountability + system reform) often leads to diluted insights. Be intentional about the primary purpose – others can be secondary.*

Linkage: *The same object can support multiple evaluations over time. A program might begin with a formative review, end with a summative evaluation, and later be included in a thematic or impact evaluation.*

4.2.5. STEP 5: SELECT THE EVALUATION TYPE

This is the synthesis point – where the object, timing, and purpose come together to determine the most appropriate evaluation type. Based on your previous responses, you will arrive at one of the following:

TABLE 24: SELECTION OF EVALUATION TYPE

Object	Timing	Use	Recommended Evaluation Type
Regulation	Ex-Ante	Design Support	Regulatory Impact Assessment (RIA)
Regulation	Ex-Post	Performance, Burden	Ex-Post RIA
Policy	Ex-Ante	Strategic Design Validation	Policy Impact Assessment (PIA)
Policy	Ex-Post	Coherence, Results	Policy Evaluation / Ex-Post PIA
Program	Mid-Term	Learning	Formative Program Evaluation
Program	Final	Accountability	Summative Program Evaluation
Multiple	Final	Joint Contribution	Cluster Evaluation
Sector	Periodic	Alignment and Reform	Sectoral or Strategic Evaluation
Any	Ex-Post	Attribution / Long-Term Impact	Impact Evaluation
System	Periodic	Evaluation System Performance	Meta-Evaluation

Tip: *Evaluation types are not mutually exclusive. A sectoral evaluation may include policy evaluations and impact assessments within it.*

4.2.6. STEP 6: CONFIRM TOOLS, STAKEHOLDERS, AND READINESS

- Before initiating the evaluation, practitioners should confirm operational readiness. This includes:
- **Methods and tools:** Will you use theory of change, cost-benefit analysis, outcome harvesting, or contribution analysis?
- **Data and evidence:** Is baseline data available? Are administrative systems reliable?
- **Stakeholders:** Who will manage the evaluation? Who must be consulted? Who will act on the findings?
- **Resources:** Is there a realistic budget? Is the evaluation aligned with the annual planning and budget cycle?

Tip: Even a well-sscoped evaluation can fail if the team hasn't checked that the timing, capacity, or data support its execution.

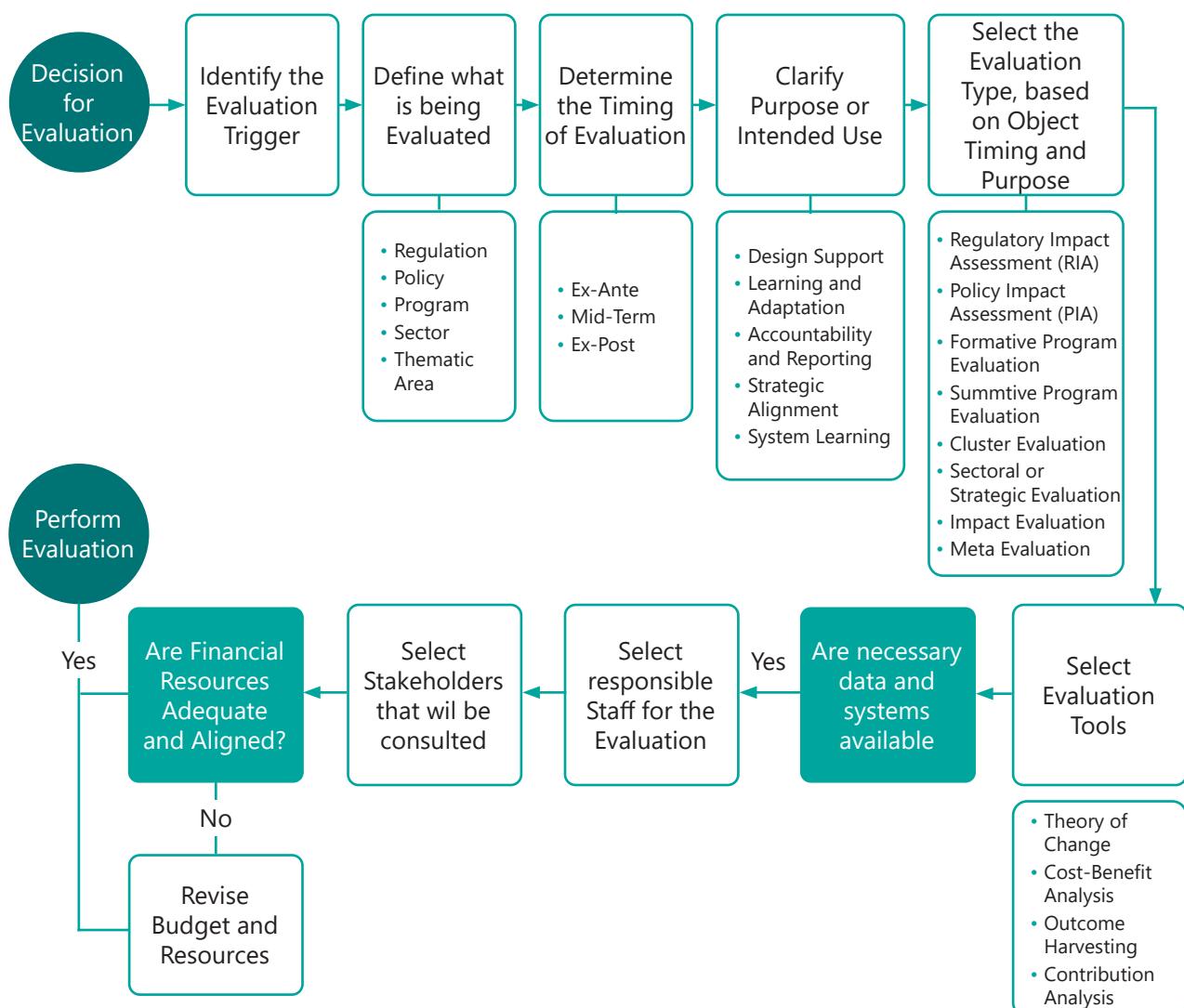


Figure 3 Process Map for Selection of Evaluation Type

5. MONITORING IN A NATIONAL M&E SYSTEM

Monitoring is the systematic and continuous process of collecting, analyzing, and using information to track the implementation of policies, programs, and regulations. It serves as a critical management tool that enables stakeholders to assess progress, identify challenges, and make informed decisions to enhance performance and accountability. Effective monitoring ensures that interventions remain aligned with their objectives, resources are utilized efficiently, and desired outcomes are achieved. It also provides the necessary data to inform evaluations, policy adjustments, and strategic planning.

5.1. TYPES AND PURPOSES OF M&E

Monitoring is not a single activity, but a set of complementary functions that operate across different domains of public sector performance. Depending on the policy or regulatory context, different types of monitoring may be required. Some forms are operational and focused on inputs or activities; others are strategic and aligned to outcomes and institutional performance. Understanding the purpose of monitoring is essential for selecting appropriate tools, assigning responsibilities, and interpreting results.

In the context of a national M&E system, eight key types of monitoring can be distinguished. Each responds to a specific information need and is typically associated with a distinct level of decision-making. Often, more than one type will apply simultaneously, particularly in large or multisectoral interventions.

- **Process and results monitoring** are the foundation of most government performance systems, allowing implementers to assess whether delivery is happening as planned and whether short-term objectives are being achieved.
- **Compliance and financial monitoring** ensure accountability and legal fidelity, often feeding into audit and reporting functions.
- **Contextual monitoring** helps track shifts in the environment that may affect implementation, while **beneficiary monitoring** emphasizes user experience, accessibility, and equity.
- **Organizational and participatory monitoring** enhance transparency, adaptive learning, and local ownership, especially where decentralization or community engagement is a feature of policy delivery

TABLE 25: MONITORING TYPOLOGIES

Monitoring Type	Purpose	Focus	Common Tools
Process Monitoring	Track implementation fidelity and operational progress	Inputs, activities, timelines	Work plans, activity trackers, supervision reports
Results Monitoring	Assess progress toward intended short- and medium-term outcomes	Output and outcome-level indicators	Indicator dashboards, performance scorecards
Compliance Monitoring	Ensure adherence to standards, regulations, and procedural obligations	Legal, regulatory, and contract compliance	Checklists, audit reports, inspection records

Monitoring Type	Purpose	Focus	Common Tools
Financial Monitoring	Track use of resources for efficiency, transparency, and budget execution	Budgets, expenditures, financial flows	Budget tracking tools, variance analysis, expenditure reports
Context Monitoring	Track changes in external conditions that may influence implementation	Political, environmental, economic or security trends	Risk scans, situation reports, scenario assessments
Beneficiary Monitoring	Capture feedback and satisfaction from service users or target groups	Experiences, accessibility, satisfaction	Surveys, citizen scorecards, grievance mechanisms
Organizational Monitoring	Evaluate internal capacity, systems, and performance of delivery units	HR systems, management processes, institutional practices	Staff evaluations, capacity assessments, internal audits
Participatory Monitoring	Promote accountability and local ownership through co-monitoring	Community-defined priorities and citizen engagement	Community scorecards, participatory appraisals, public hearings

In practice, the boundaries between these types are not rigid. For example, a well-designed beneficiary monitoring system may yield compliance insights, while organizational monitoring may surface data relevant to strategic outcomes. What matters most is that the monitoring system is aligned to its **intended use**, and that data collected through each channel is ultimately used for learning, correction, and improved performance.

In the chapters that follow, several of these monitoring types are embedded within evaluation frameworks, sectoral M&E strategies, and digital reporting platforms. Each type plays a role in establishing an integrated, performance-oriented public administration.

5.2. MONITORING ACROSS THE RESULT CHAIN

Monitoring must be grounded in the logical structure of the intervention it is tracking. This structure – often visualized as a **results chain** or **results framework** – articulates how inputs (resources) are converted into activities, how these generate outputs (products or services), and how these outputs are expected to lead to outcomes (behavioral or institutional change), and ultimately to long-term impacts.

Each level of the results chain requires a different type of monitoring focus, indicator design, data source, and frequency. Monitoring without reference to this structure risks measuring what is available rather than what is useful.

For public sector actors, the practical value of aligning monitoring to the results chain is threefold:

- It clarifies **who is responsible** for which results (e.g., outputs at program level, outcomes at strategic or inter-agency level).
- It links **monitoring efforts to policy decision cycles**, from operational troubleshooting to mid-term review and strategic redirection.

- It ensures that **monitoring data feeds meaningfully into evaluation** – providing the baseline, trend data, and early performance signals needed to determine evaluation timing, focus, and scope.

In most systems, monitoring is strongest up to the output level, where data is easier to collect through administrative channels. Monitoring of outcomes requires more deliberate design and often overlaps with evaluation, as outcomes may take time to materialize and are harder to directly attribute.

Below is a summary of the core elements of the results chain and how each is typically monitored.

TABLE 26: MONITORING ACROSS THE RESULT CHAIN

Results Level	Definition	Monitoring Focus	Typical Tools & Data Sources
Inputs	Resources mobilized to support an intervention	Budget execution, staffing, equipment availability	Budget tracking tools, procurement reports, HR records
Activities	Actions undertaken using inputs	Timeliness, quantity, and quality of service delivery	Work plans, activity reports, supervision checklists
Outputs	Direct products or services delivered as a result of activities	Delivery rates, access levels, service coverage	MIS, program statistics, service delivery scorecards
Outcomes	Short- to medium-term effects of outputs on beneficiaries or systems	Behavior change, institutional uptake, service utilization	Surveys, routine administrative data, outcome dashboards
Impacts	Long-term changes in society, the economy, or the environment	Poverty, employment, health, education, resilience	National surveys, sector reports, impact assessments

5.3. MONITORING FUNCTIONS AND RESPONSIBILITIES

A monitoring system is only effective when its responsibilities are clearly defined and distributed across institutions. Monitoring is not the task of a single actor – it is a shared function, embedded within planning, budgeting, implementation, and oversight processes. However, without clear roles and institutional ownership, monitoring risks becoming fragmented, duplicative, or simply neglected.

Each level of government and institutional actor plays a different but complementary role in the monitoring chain. These roles vary by:

- Level of responsibility (national, sectoral, regional, program-specific),
- Type of intervention (policy, regulation, program, or budget),
- Use of monitoring outputs (management, accountability, learning, or compliance).

In a robust M&E system, there are four core dimensions to institutional monitoring responsibilities:

- 1. Data Generation:** Institutions collect and compile data relevant to their functions.
- 2. Data Verification and Quality Control:** Mechanisms are in place to check accuracy, consistency, and integrity of reporting.
- 3. Data Analysis and Interpretation:** Monitoring data is not only reported but interpreted to support learning and decisions.

4. Reporting and Feedback Loops: Monitoring results are fed back into planning, reform, and policy cycles – closing the loop between evidence and action.

Importantly, this system must allow for both vertical accountability (upward reporting to central authorities and donors) and horizontal learning (cross-sectoral or regional comparison and adaptation).

TABLE 27: STAKEHOLDER ROLES IN MONITORING

Institution	Primary Monitoring Functions
Line Ministries / Implementing Agencies	Track implementation of policies and programs under their mandate; report on inputs, activities, and outputs.
Sectoral M&E Units	Coordinate results monitoring across sectoral interventions; oversee indicator frameworks; ensure data consistency.
Local Governments / Regional Authorities	Collect and report on localized data; monitor access and quality of service delivery; relay feedback from users.
Central M&E Unit / Coordination Body	Consolidate performance data across sectors; develop standards, templates, and guidelines; link M&E to strategic planning and budgeting.
National Statistics Office / Technical Agencies	Ensure statistical integrity and alignment between monitoring data and official statistics; advise on indicator design.
Audit and Oversight Bodies	Verify compliance with reporting standards; review data accuracy; use monitoring data for budget and legal oversight.
Development Partners / Donors	Provide technical assistance and funding for monitoring frameworks; may require sector- or program-level performance reporting.
Civil Society Organizations / Media	Monitor delivery on the ground; represent citizen voice; conduct independent tracking and social accountability.

Institutional Considerations

When designing or strengthening a national monitoring system, several institutional considerations must be addressed:

- **Mandate clarity:** Ensure that every institution knows what it must monitor and for whom. Avoid overlapping roles.
- **Capacity development:** Equip staff with the tools, training, and systems needed to collect and analyze data.
- **Incentives for use:** Create institutional demand for monitoring data by linking it to planning, budgeting, or decision-making.
- **Technology and interoperability:** Build digital systems (e.g., dashboards, MIS) that allow institutions to exchange information easily and in real time.
- **Legal and procedural alignment:** Embed monitoring roles into legal frameworks, decrees, and reporting templates – so that they become institutionalized and not ad hoc.

5.4. MONITORING TOOLS AND MODALITIES

Effective monitoring relies on more than just a list of indicators – it requires a set of structured tools and platforms to support consistent data collection, analysis, and reporting across all levels of government. These tools serve different functions: some are used to organize and

plan monitoring efforts, others to generate or visualize data, and others still to translate that data into insight and decision-making.

The selection of tools must reflect:

- The type and scale of the intervention (e.g., program, regulation, policy);
- The results level being tracked (e.g., output vs. outcome);
- The level of formality and frequency of reporting required;
- The degree of digitalization and interoperability across institutions;
- The users of monitoring data (e.g., program managers, policy planners, oversight bodies).

Institutionalization of tools is critical: without standard templates, data systems, and feedback formats, even well-designed M&E frameworks can struggle to generate actionable intelligence. Conversely, tools that are overly complex, underfunded, or misaligned with local capacity risk being underutilized or abandoned.

Many of the tools used for monitoring are also relevant to evaluation and planning, reinforcing the importance of using integrated systems that minimize duplication and maximize learning across the public sector.

TABLE 28: MONITORING TOOLS

Tool / Platform	Function	Typical Users
Monitoring Framework	Defines what will be monitored, using which indicators, at what frequency, and by whom	Program managers, M&E units
Indicator Matrix	Organizes indicators by results level, data source, disaggregation, and baseline/target	Policy teams, line ministries
Performance Dashboards	Visualize key results for decision-makers; often used in real-time settings (e.g., Cabinet briefings)	High-level officials, central agencies
Management Information Systems (MIS)	Store and manage large volumes of routine monitoring data, especially at activity and output levels	Implementation agencies, sector ministries
Progress Reports	Provide structured summaries of delivery progress, often on quarterly or semi-annual cycles	Implementing units, donors
Monitoring Scorecards	Aggregate service-level results (e.g., school performance, district health delivery) in comparative formats	Sector leads, regional governments
Field Supervision Checklists	Used during on-site visits to verify implementation and assess adherence to plans	Technical teams, inspection bodies
Citizen Feedback Tools	Capture user perspectives through participatory monitoring (e.g., report cards, grievance platforms)	Civil society, advocacy organizations
Mobile and Digital Data Collection	Enables field-level data collection using tablets or smartphones; linked to MIS platforms	Enumerators, regional M&E teams

6. A CONCISE GUIDE TO REGULATORY IMPACT ASSESSMENT (EX-POST RIA)

6.1. WHAT IS EX-POST RIA AND WHY IT MATTERS

Ex-Post Regulatory Impact Assessment (RIA) is a structured, evidence-based approach used to evaluate the effects of a regulation after it has been implemented. Ex-post RIA focuses on understanding **whether a regulation achieved its intended objectives, whether it generated any unintended consequences**, and how it **performed in terms of efficiency, proportionality, and public value**.

Unlike ex-ante RIA, which assesses the likely impacts of a regulation before it is adopted, ex-post RIA is **retrospective**. It examines how a regulation has operated **in practice** analysing **compliance behaviour, implementation dynamics, and observable outcomes**. The goal is not only to review performance, but also to **identify whether a regulation should be amended, replaced, repealed, or scaled**.

While ex-post RIA shares broad evaluation principles with ex-post Policy Impact Assessment (PIA), it applies **specifically to legally binding normative acts**. The focus is typically narrower, often centred on a **defined regulatory obligation** or **compliance mechanism**, and it uses specialized tools such as **administrative burden assessments, legal coherence checks, and post-implementation reviews** of regulatory effectiveness. Throughout the guide, the methodology will be complemented by an example, based on the 2020 IT Park Residency Regulation. This refers to a fictional evaluation of the real IT Park residency regulation.

In Uzbekistan's digital transformation journey, ex-post RIA serves several essential functions:

- **Assessing the real-world impact of flagship reforms:** Regulations such as the **IT Park Residency** or **digital services licensing reforms** have been central to the government's ambitions to build a robust innovation ecosystem. However, beyond formal adoption, ex-post RIA allows evaluators over the conversation from theoretical potential to tangible evidence of performance by asking to ask: Have these reforms **actually spurred digital entrepreneurship**, innovation, or foreign investment? Did the **incentives reach the intended beneficiaries**, or were they primarily captured by a narrow set of actors? Are the **costs to the public budget justified** by the resulting outcomes?
- **Testing whether expected outcomes materialized:** The regulation may have been well-intentioned and implemented, but did it deliver? Ex-post RIA provides a mechanism to track whether:
 - Startup registration and survival rates actually improved post-regulation;
 - Youth employment in the ICT sector increased as projected;
 - Export performance of resident digital firms showed meaningful growth.
- **Uncovering bottlenecks and unintended effects:** Even well-designed policies can produce unexpected or undesirable outcomes. For instance female and youth entrepreneurs may have been unintentionally excluded due to unresponsive eligibility criteria or lack of outreach. Ex-post RIA helps flag such dynamics early, enabling **course correction** before institutional inefficiencies become entrenched.
- **Informing revisions to regulatory frameworks:** Uzbekistan's digital economy is not static – it is evolving rapidly in response to market shifts, global investment trends,

and technological change (e.g., AI, blockchain, e-commerce). By systematically capturing what worked and what didn't, ex-post RIA:

- Enables regulatory fine-tuning or sunset provisions;
- Provides feedback loops to improve future legislation;
- Aligns reforms with international standards and competitiveness benchmarks (e.g., OECD digital government indicators, Global Innovation Index).

6.2. A STEP-BY-STEP GUIDE TO CONDUCTING EX-POST RIA

STEP 1. DEFINING THE REGULATORY OBJECT AND PURPOSE OF THE RIA

The first step in conducting a RIA is defining the object, and its purpose. The object of an ex-post RIA is a **regulation that has already been in force for a defined period** – such as a government resolution, ministerial order, or special decree. The aim is to evaluate:

- Whether the regulation addressed the problem it was designed to solve;
- Whether compliance occurred as expected and produced public value;
- Whether the regulation remains relevant and proportionate in its current form;
- And whether it should be continued, amended, replaced, or withdrawn.

This initial step involves:

- Clarifying the legal status, scope, and intent of the regulation;
- Reviewing original policy documents, cabinet notes, or regulatory justifications;
- Defining the evaluation timeframe and geographic or sectoral coverage;
- Identifying who commissioned the ex-post RIA, and for what purpose (e.g. legal review, simplification initiative, reform planning).

STEP 1 EXAMPLE: IT PARK RESIDENCY REGULATION

*In this case, the object of evaluation is the **IT Park Residency Regulation**, amended by **Resolution No. 589 (2020)**. The IT Park residency regulation established eligibility criteria and tax exemptions for ICT firms operating within designated technology zones. After three years of implementation, a decision was made to conduct an ex-post RIA to evaluate whether the regulation had effectively attracted digital enterprises, stimulated investment, and reduced informality. The review aimed to inform whether the residency provisions were functioning as intended or required revision. This regulation is selected because:*

- **It is a core pillar of Digital Uzbekistan 2030.**
- **It represents a major fiscal and economic intervention** in the tech/startup sector.
- **It has nationwide implementation implications**, including equity and regional development aspects.
- **Enough time (2020–2024) has elapsed to assess early performance and emerging impacts.**

STEP 2. DEFINING THE PURPOSE AND SCOPE OF THE EVALUATION

After identifying the regulation to be assessed, the next step in the ex-post RIA process is to clearly define the **purpose and scope** of the evaluation. This step sets the direction of the entire

assessment and ensures that efforts are aligned with the specific policy questions that matter most to stakeholders. The **purpose** of an ex-post RIA can vary depending on the regulatory context and policy priorities – it may aim to verify whether the regulation achieved its intended objectives, identify implementation challenges or bottlenecks, evaluate cost-effectiveness, or examine unintended or distributional impacts, such as gender or regional disparities.

Once the purpose is defined, evaluators must then determine the **scope** of the evaluation. This involves making practical choices about what to include and what to exclude in order to maintain analytical focus and feasibility. The scope may cover the **entire regulation or specific provisions** that are of particular interest – such as tax incentives, licensing rules, or compliance mechanisms. It may focus on the **national level** or select **specific regions or implementation sites**, especially if regional disparities in delivery or outcomes are suspected. Similarly, the evaluation might aim to assess **comprehensive outcomes** across all dimensions of the regulation or limit itself to a **particular area of interest**, such as youth entrepreneurship, digital exports, or support for female-led firms.

These decisions require careful judgment. While a broad scope may offer a more complete picture, it also increases the complexity, data demands, and resource requirements of the evaluation. A **narrower, well-targeted scope**, on the other hand, allows for deeper analysis and more **actionable recommendations**, especially when the objective is to inform revisions to specific regulatory mechanisms or to scale up successful components. Balancing **comprehensiveness with feasibility** is therefore essential to ensuring that the evaluation generates insights that are both rigorous and usable. The scope of the evaluation, is closely linked to its type, as mentioned in previous sections.

STEP 2 EXAMPLE: PURPOSE AND SCOPE OF THE IT PARK REGULATION

Purpose:

To assess whether the IT Park residency regime has effectively promoted digital entrepreneurship, regional development, and startup growth, in line with its intended objectives under the national digitalization strategy.

Scope:

- Focus on **three specific policy levers** within the regulation:
 1. **Tax incentives and exemptions**
 2. **Startup registration and certification process**
 3. **Regional extension of IT Park hubs beyond Tashkent**
- Thematic focus on **startup survival, geographic equity, and female entrepreneur inclusion**

STEP 3. RECONSTRUCTING THE INTERVENTION LOGIC

Ex-post RIA begins by **reconstructing the logic that underpinned the regulation at the time of its adoption**. This involves examining the original rationale, what problem the regulation was meant to address, what behavior it aimed to influence, and what results it expected to produce.

The evaluator should:

- Identify the **problem definition** in the original policy narrative;
- Document the **regulatory mechanism** (e.g., incentives, obligations, prohibitions);

- Define the **expected behavioral and systemic outcomes**;
- Surface the **assumptions** made about compliance, capacity, and cost-benefit trade-offs.

The **intervention logic** (or theory of change) articulates the **intended causal pathway** from regulatory inputs to final impacts. This model helps structure the evaluation.

Components typically include.



Figure 4: Intervention Logic

- **Inputs** – resources, regulatory instruments, funding
- **Activities** – administrative actions, enforcement, outreach
- **Outputs** – number of licenses, services delivered
- **Outcomes** – changes in stakeholder behavior or performance
- **Impacts** – broader economic, social, or environmental changes

This model ensures that evaluation criteria and indicators are clearly linked to the original intentions of the regulation.

STEP 3 EXAMPLE: INTERVENTION LOGIC

The intervention logic based on the IT regulation example is as follows:

Level	IT Park Regulation Example
Inputs	Budgetary support for IT Park operations; legal framework; staff
Activities	Startup outreach, application processing, certification
Outputs	# of registered resident firms; # of firms receiving tax exemptions
Outcomes	Increased startup survival; expansion to regional hubs
Impacts	Digital sector job creation; economic diversification; innovation

STEP 4: STAKEHOLDER MAPPING AND INSTITUTIONAL LANDSCAPE

A policy is rarely implemented by a single actor. Ex-post PIA requires mapping out the full network of institutions and stakeholders involved in policy design, coordination, and delivery.

The evaluator should identify:

- **Policy owner** (e.g. line ministry or inter-agency task force);
- **Implementation partners**, including agencies, regional governments, and service providers;
- **Target beneficiaries**, such as firms, citizens, or institutions;
- **Support actors**, including donors, private partners, and civil society organizations.

The mapping exercise supports later steps by clarifying who has responsibility for results, who has influence over delivery, and where coordination breakdowns may have occurred.

STEP 4 EXAMPLE: KEY STAKEHOLDER MAPPING

Key actors included the **Ministry of Digital Technologies** (policy owner), **IT Park Uzbekistan** (implementation lead), and the **Ministry of Economy and Finance** (responsible for tax exemptions). Universities hosted innovation hubs in partnership with local khokimiyats. Evaluators documented institutional coordination challenges, especially in regions where tech infrastructure and staffing capacity were weaker.

STEP 5. SETTING EVALUATION QUESTIONS AND CRITERIA

A critical component of ex-post RIA is the formulation of **clear, targeted evaluation questions** that guide the analysis and structure the findings. These questions should reflect the **priorities of policymakers**, the **concerns of affected stakeholders**, and the **underlying logic of the regulation itself**. To ensure rigor and comparability, it is common practice to align evaluation questions with **internationally recognized evaluation criteria** (e.g. OECD-DAC), which provide a comprehensive framework for assessing public interventions. Each criterion addresses a distinct dimension of regulatory performance – from the extent to which objectives were achieved, to how efficiently resources were used, and whether the regulation has generated sustainable and inclusive outcomes. These criteria help evaluators move beyond narrow compliance checks and instead generate deeper insights into the regulation's design, implementation, and real-world effects.

In an ex-post RIA, selecting the appropriate type of evaluation is crucial for ensuring that the assessment is both methodologically sound and proportionate to the regulatory context. There are three main types of evaluation, each addressing a different layer of policy analysis. **Process Evaluation** examines whether the regulation was implemented as intended – looking at procedural aspects such as institutional roles, timelines, outreach, and compliance mechanisms. This is often the starting point, as no meaningful performance or impact evaluation can be conducted without first confirming that the regulation was actually delivered as designed.

The evaluation team should:

- Define **criteria**, typically based on the OECD DAC framework: relevance, effectiveness, efficiency, impact, sustainability, and coherence;
- Formulate **evaluation questions** that reflect policy objectives;
- Select appropriate **methods**:
 - Theory-based approaches (e.g., contribution analysis);
 - Outcome tracking or before/after comparisons;
 - Qualitative methods (e.g., outcome harvesting, stakeholder interviews);
 - Surveys for beneficiary-level results or equity assessment;
- Establish a **disaggregation plan** to capture gender, geographic, and demographic dimensions.

STEP 5 EXAMPLE: SETTING EVALUATION QUESTIONS

The evaluation applied five DAC criteria. Methods included a structured beneficiary survey contribution analysis (linking service access to outcomes), and outcome harvesting (documenting spin-offs and mentorship effects). Gender and urban/rural gaps were analyzed through cross-tabulations of survey data and IT Park registries.

Uzbekistan Application: Key Questions

- Was the IT Park residency process **transparent and accessible across all regions?**
- To what extent did **tax incentives** influence startup formation and survival?
- Did the regulation result in **regional disparities or concentration in Tashkent?**
- Were **female-led startups** equally likely to register and benefit?

What is the **cost to the public budget** per surviving startup, and is it justified?

STEP 6. DATA COLLECTION AND TRIANGULATION

A robust ex-post Policy Impact Assessment relies on well-planned, multi-sourced data collection. Because public policies tend to operate across institutions and populations and often lack a single, central monitoring mechanism evaluator must piece together a **composite picture of performance** using multiple, complementary forms of evidence. This approach enables triangulation, which strengthens credibility, balances institutional bias, and helps fill gaps in official data.

Key data sources and techniques include:

- **Administrative data:** Collected by government agencies and implementing bodies (e.g., beneficiary registries, tax filings, subsidy allocations, budget disbursements). This data can show policy reach, uptake, and procedural efficiency.
- **Monitoring and internal reporting:** Often generated during policy implementation (e.g., dashboards, mid-term reviews, internal memos). These are useful for reconstructing implementation fidelity and institutional intent.
- **Surveys and questionnaires:** Deployed to reach beneficiaries, non-beneficiaries, front-line implementers, or service providers. These offer insight into behavior, satisfaction, bottlenecks, and outcome variation.
- **Key informant interviews and field visits:** Capture institutional memory, stakeholder perspectives, and contextual nuance – especially when official records are incomplete or contradictory.
- **External and third-party sources:** Useful for benchmarking, verification, or contextualization (e.g., national statistics, donor assessments, independent studies).

Disaggregation is critical for ensuring that the policy's effects on different subgroups (e.g., women, youth, rural communities, sector-specific enterprises) are fully understood. When disaggregated administrative data is unavailable, evaluators should incorporate targeted sampling and qualitative methods to uncover equity issues.

STEP 6 EXAMPLE: DATA COLLECTION AND TRIANGULATION

To evaluate the Startup Ecosystem Support Policy (2019–2023), the evaluation team applied a mixed-methods data collection plan designed around the core policy components.

- From **IT Park's administrative system**, the team retrieved firm-level data on registrations, sector focus, region, and survival status (2 years post-residency).
- The **Ministry of Economy and Finance** provided tax filing records and export declarations for startups that accessed the fiscal incentive scheme.

- To assess user-level experiences and inclusion, a **structured survey** was deployed to 312 startups across six regions. The instrument included questions on access to services, mentoring quality, gender dynamics, and growth constraints.
- In addition, **22 semi-structured interviews** were conducted with startup founders, innovation hub coordinators, and private accelerators.
- In regions where activity was low or unreported, the evaluation team conducted **site visits** to document contextual barriers – including infrastructure, staff turnover, and outreach limitations.
- Finally, **external sources** such as the World Bank's digital economy diagnostics and UNDP's regional competitiveness studies were consulted to benchmark national progress.

STEP 7. ANALYSING AND INTERPRETING FINDINGS

Once data has been collected from multiple sources, the next step is to **analyse and interpret the findings** through the lens of the regulation's original intervention logic. This means assessing whether the expected results from outputs to long-term impacts actually materialized, and if not, understanding why. A key principle in this phase is **triangulation**: using different types of data and perspectives to cross-check conclusions and reduce the risk of bias or over-reliance on a single data stream.

Several methods can support this analytical process. **Gap analysis** involves comparing what was planned with what actually occurred, identifying areas where implementation fell short or exceeded expectations. **Contribution analysis** is particularly useful when multiple factors may have influenced outcomes; it helps determine the extent to which observed changes can reasonably be attributed to the regulation, as opposed to other programs or external trends. In contexts like Uzbekistan, where donor interventions often overlap with government initiatives, this method ensures that conclusions are appropriately nuanced. Lastly, **comparative analysis** can highlight spatial or demographic disparities by examining differences between regions, between male- and female-led firms, or between beneficiaries and non-beneficiaries of the policy.

STEP 7 EXAMPLE: ANALYSING AND INTERPRETING FINDINGS

In the case of the IT Park Residency Regulation, an ex-post RIA could potentially uncover several important trends warranting further exploration. For example, it may identify implementation gaps across regional IT Parks, particularly related to administrative capacity or procedural efficiency. Some hubs might show signs of delayed processing of residency applications or inconsistent application of eligibility criteria, potentially discouraging firm participation outside the capital. The evaluation could also highlight disparities in startup survival rates, with firms in Tashkent possibly performing better than those in other regions – suggesting unequal access to support mechanisms such as mentoring, investor networks, or digital infrastructure. Moreover, the RIA might reveal gender-based disparities, such as low representation of women-led startups among certified residents, raising questions about the inclusivity and reach of the regulation's implementation. Collectively, such findings would help pinpoint both outcome-level differences and underlying structural barriers, providing a critical evidence base for improving the equity, effectiveness, and territorial balance of future regulatory measures.

STEP 8. REPORTING, VALIDATION AND RECOMMENDATIONS

The final stage of the ex-post RIA process involves consolidating findings into a **clear, structured, and policy-relevant report**. The report should include all key components – such as an executive summary, a discussion of findings, and a set of evidence-based recommendations – presented in a way that is **accessible to both technical experts and non-specialist decision-makers**. The aim is not only to communicate what the regulation achieved, but also to highlight where improvements are needed and how future interventions can be made more effective. Importantly, the reporting process should include a **validation phase**, in which preliminary findings are shared with key stakeholders – such as implementing agencies, beneficiaries, and policymakers – to gather feedback, test assumptions, and build consensus around next steps. In addition to the full technical report, the evaluation team may prepare a **dashboard of performance indicators** for use by monitoring bodies, and **concise policy briefs** that translate results into **actionable insights** for decision-makers. These deliverables ensure that the findings are not only documented, but actively used to inform policy revision and future regulatory design.

Findings must be synthesized into a form that can inform policy adjustments and strategic decisions. The evaluation should:

- Prepare a **summary report** with actionable recommendations;
- Provide a **technical annex** detailing data sources, methods, and limitations;
- Facilitate **validation workshops** with key institutions;
- Develop a **management response** or policy adjustment plan based on conclusions.

The final PIA product should support planning, budgeting, or reform of the next policy phase.

STEP 8 EXAMPLE: REPORTING, VALIDATION AND RECOMMENDATIONS

The outputs of an ex-post RIA should be tailored to meet the needs of diverse stakeholders, ensuring that findings are both actionable and widely disseminated. Key output formats may include:

Evaluation Report submitted: A comprehensive technical report detailing the full scope of the evaluation – covering methodology, data sources, findings, and policy recommendations.

Policy Brief for MEF, including a budget-cost analysis of startup tax relief: A concise, targeted brief would be prepared for the Ministry of Economy and Finance, highlighting the key economic and fiscal findings of the evaluation. It would include a budgetary analysis of the cost-effectiveness of tax exemptions granted to IT Park residents and their contribution to policy objectives like startup survival or job creation.

Validation Workshop with regional IT Park coordinators, startup founders, and donor partners: A participatory validation event would bring together key stakeholders to review the preliminary findings, reflect on implementation experiences, and co-develop a set of practical recommendations.

7. A CONCISE GUIDE TO POLICY IMPACT ASSESSMENT

7.1. INTRODUCTION TO EX-POST POLICY IMPACT ASSESSMENT (PIA)

Ex-post Policy Impact Assessment (PIA) is a structured evaluation method used to assess the effects of a public policy **after it has been implemented**. Its primary purpose is to determine whether **a policy achieved its intended results**, understand how and why those results were (or were not) realized, and generate actionable learning to inform future decision-making. While project evaluations often focus on operational delivery and short-term outputs, ex-post PIAs examine **broader public policies**, which tend to span institutions, sectors, and delivery mechanisms – and often pursue ambitious, strategic transformations.

In Uzbekistan, where cross-sectoral initiatives are central to national development planning, ex-post PIA is emerging as an essential tool for **policy accountability, reform readiness, and strategic foresight**. Many of the policy instruments under the **Digital Uzbekistan 2030 Strategy** exemplify the kind of interventions that benefit from this approach. They involve multi-stakeholder coordination, systemic change goals, and public–private partnerships that cannot be fully understood through monitoring or program reviews alone. ex-post PIA offers a way to look beyond operational delivery and assess deeper **systemic outcomes**. It helps answer critical questions like:

- Did the policy solve the problem it was designed to address?
- Were the benefits distributed equitably?
- How did implementation unfold across institutions and regions?
- What should be adjusted, scaled, or sunset in future phases?

7.1.1. WHY A POLICY-CENTRIC EVALUATION APPROACH?

Public policies are not limited to individual programs or regulations they operate as systems of coordinated effort, often intended to influence long-term societal or economic trends. This complexity means that **policy evaluations must go beyond simply “checking if targets were met”**. They must examine the **policy’s logic, assumptions, design robustness, and responsiveness to real-world conditions**. Ex-post PIA allows evaluators to assess **whether a policy truly solved the problem it was designed to address**, and whether its effects were equitable, sustainable, and strategically coherent.

A well-executed ex-post PIA can:

- Determine the **effectiveness** of a policy in achieving intended outcomes.
- Identify **unintended or distributional effects**, particularly among vulnerable populations.
- Assess **institutional and stakeholder coherence** across implementation actors.
- Contribute to **policy refinement, replication, or decommissioning**.
- Provide evidence to inform **new strategic choices or investment priorities**.

Example Introduction: To illustrate the ex-post PIA process throughout this section, we will use as a reference a mock initiative of the Government of Uzbekistan under **Digital Uzbekistan 2030: the ICT Startup Ecosystem and Innovation Cluster Support Program**.

The mock program is to have been launched in 2019 and had the following objectives:

- Establish and expand **IT Parks** and **regional innovation clusters**;
- Provide **startup incubation, mentorship, and access to seed capital**;
- Offer **tax and regulatory incentives** for digital entrepreneurs;
- Encourage **technology transfer, R&D linkages, and internationalization** of Uzbek digital firms;
- Create **university-based incubation centers** and promote youth engagement in tech.

By the end of 2023, the policy had entered its final implementation phase, with over a dozen IT Parks established, thousands of jobs created, and increased visibility of Uzbekistan's tech sector in regional innovation rankings. However, decision-makers raised questions about the **depth and sustainability** of impact, especially for startups outside of Tashkent, the long-term survival of incubated firms, and whether the policy had genuinely addressed barriers for women and rural entrepreneurs.

Throughout this chapter, we will refer back to this fictional case to demonstrate how ex-post PIA works in practice, illustrating methodological choices, evaluation criteria, stakeholder engagement, and analytical tools at each stage of the process.

7.1.2. THE LOGIC OF EX-POST POLICY IMPACT ASSESSMENT (PIA)

Evaluating a public policy after it has been implemented is a fundamentally different exercise than monitoring programs or conducting compliance checks. Ex-post Policy Impact Assessment (PIA) **demands a return to the original policy intent**, an interrogation of what change was meant to occur and for whom, and a careful analysis of what actually happened – intended or not. This requires both a conceptual framework and a set of analytical principles capable of unpacking complex, real-world interventions.

FROM POLICY OBJECTIVES TO POLICY EFFECTS

Every policy, whether explicitly stated or not, is built upon a **theory of change**: a set of assumptions about how a given problem will be solved through specific actions. In ex-post PIA, one of the first tasks is to **reconstruct this theory**: what outcomes were envisioned, what mechanisms were expected to generate them, and under what conditions they were meant to occur.

This backward-facing analysis does not simply test if outputs were delivered – it asks:

- Were the original assumptions valid?
- Did causal pathways play out as expected?
- Were certain groups left behind?
- Did contextual shifts affect outcomes?
- Were there unintended effects?

Ex-post PIA also allows evaluators to **look beyond delivery metrics**, to explore questions of policy design quality, relevance in a changed context, and systemic coherence.

EXAMPLE STARTUP ECOSYSTEM POLICY:

The policy anticipated that establishing regional tech parks and incubators would lead to startup growth, innovation, and job creation. But did tech entrepreneurs in smaller cities actually access these services? Did they receive investment and mentoring? Did they remain viable after exiting incubation? Were women-led startups equally represented? Ex-post PIA seeks to answer such questions with evidence, not assumptions.

APPLYING THE OECD DAC CRITERIA AT THE POLICY LEVEL

While originally developed for international development evaluations, the OECD DAC criteria are widely used in policy evaluation and offer a flexible but structured lens for ex-post PIA:

- **Relevance:** Was the policy well-aligned with the problem it sought to solve? Is it still relevant today?
- **Effectiveness:** Did the policy achieve its intended results? To what extent and for whom?
- **Efficiency:** Were the resources used proportionate to the benefits achieved?
- **Impact:** What broader, longer-term changes occurred? Were there unintended outcomes?
- **Sustainability:** Will results continue without the policy's continued support?
- **Coherence (in newer frameworks):** Did the policy complement or conflict with other policies or programs?

Evaluators may prioritize some criteria over others depending on the scope and intended use of the assessment.

EXAMPLE POLICY COHERENCE QUESTION:

Did Uzbekistan's ICT startup policy work in harmony with higher education reforms, foreign investment promotion, and national employment strategies or did fragmentation between ministries undermine its system-wide potential?

7.1.3. CONTRIBUTION OVER ATTRIBUTION

Most public policies—especially complex, innovation-driven ones—do not operate in isolation. For this reason, ex-post PIA often moves beyond trying to “**attribute**” change solely to one policy. Instead, it examines how the policy **contributed** to change, in interaction with other forces (e.g., private investment, market trends, international developments).

This shift calls for **theory-based approaches** like:

- **Contribution analysis**
- **Outcome harvesting**
- **Process tracing**
- **Realist evaluation**

These allow for credible and transparent assessments of how and why observed results came about, even in the absence of counterfactual evidence.

SEEING THE SYSTEM, NOT JUST THE INTERVENTION

Policies often interact with institutional systems – ministries, agencies, legal frameworks, budgets. Ex-post PIA must therefore consider:

- **Institutional coherence and leadership dynamics;**
- **Implementation arrangements and decentralization;**
- **Stakeholder engagement and legitimacy;**
- **Policy adaptability in the face of uncertainty.**
- **Not Just for Judgement – Also for Learning**

Finally, the logic of ex-post PIA must balance two roles:

- **Accountability** – Did the policy deliver?
- **Learning** – What needs to change next time?

By surfacing both achievements and limitations, ex-post PIA becomes an essential tool not just for post-mortems, but for informed iteration and strategic renewal.

Insight: A well-conducted PIA should lead directly to decisions: Should the policy be extended, revised, scaled, decentralized, integrated, or replaced?

7.2. STEP-BY-STEP GUIDE TO CONDUCTING AN EX-POST PIA

7.2.1. USING THE STARTUP ECOSYSTEM SUPPORT POLICY AS AN ILLUSTRATIVE EXAMPLE

Conducting an ex-post PIA requires a systematic, yet flexible process tailored to the characteristics of the policy under review. While the specific sequence may vary depending on context, the six stages outlined below represent best practice for planning and executing a comprehensive, insight-driven policy evaluation. Each stage is illustrated through the example of Uzbekistan’s Startup Ecosystem and Innovation Cluster Support Policy under Digital Uzbekistan 2030.

STEP 1: DEFINING THE POLICY OBJECT AND PURPOSE OF THE PIA

The object of a PIA is a policy, understood as a set of coordinated interventions with a shared strategic goal – typically articulated in the form of a government strategy, reform program, or inter-ministerial initiative. A PIA may focus on:

- A sectoral policy, such as an innovation strategy or digital skills roadmap;
- A national development program, such as an employment or inclusion framework;
- A cross-cutting theme, such as gender, climate, or digital transformation.

The first step in a PIA is to define what portion of the policy will be evaluated. This includes clarifying:

- The intended outcomes of the policy, as stated at inception;
- The duration and territorial coverage of implementation;

- The implementing institutions and governance model;
- The target population or system the policy aimed to influence.

EXAMPLE STEP 1: DEFINING THE POLICY OBJECT AND PURPOSE OF PIA

Under the Digital Uzbekistan Strategy, the government supported the development of tech entrepreneurship through startup incubators, tax incentives, university-linked innovation hubs, and export facilitation measures. A PIA was commissioned after the first phase (2019–2023) to assess whether this policy had improved startup survival, enhanced access to finance, and contributed to regional innovation growth.

Unlike a RIA, which would focus on a specific regulation (such as the legal criteria for IT Park residency), the PIA evaluates the strategic ensemble – including non-regulatory activities, inter-agency collaboration, and support instruments whose effectiveness depends on institutional learning and stakeholder interaction.

STEP 2: RECONSTRUCTING THE LOGIC OF THE POLICY

Every public policy is grounded in a logic of intervention – an underlying narrative that links a problem to a strategic response and outlines how that response is expected to create change. In ex-post PIA, this logic is not taken at face value: it must be critically reconstructed to assess whether the policy unfolded as intended, whether its assumptions held, and how its design translated into action.

Reconstructing the policy logic typically involves:

- Clarifying the problem definition and what the policy sought to change;
- Identifying the intervention components, including services, incentives, partnerships, and investments;
- Mapping the expected causal chain, from inputs and activities through outputs, outcomes, and long-term impacts;
- Documenting assumptions and contextual conditions (e.g., institutional capacity, market dynamics);
- Testing for alignment between stated objectives and actual implementation paths.

EXAMPLE STEP 2: RECONSTRUCTING THE LOGIC OF THE POLICY

The Digital Entrepreneurship Policy aimed to stimulate innovation and competitiveness by expanding startup support infrastructure, offering tax relief to IT Park residents, and facilitating early-stage financing. Its logic assumed that increased access to incubation and investment would drive firm formation, boost exports, and diversify the economy. A reconstructed theory of change revealed that while infrastructure and incentives were delivered, weaker assumptions around market readiness and investor networks led to mixed outcomes.

STEP 3: MAPPING STAKEHOLDERS AND POLICY INSTITUTIONS

Unlike regulations, which are usually implemented through centralized administrative systems, policies often span **multiple institutions, levels of government, and non-state**

actors. Understanding the governance model of a policy is essential for both interpreting results and assigning responsibility.

Stakeholder mapping in PIA helps to:

- Identify the **policy owner(s)** and their coordination roles;
- Document the **network of implementers** (e.g., ministries, agencies, private partners, local authorities);
- Capture the **beneficiary spectrum** (by gender, age, sector, region);
- Surface **interdependencies and bottlenecks**, including gaps in mandate clarity or overlapping initiatives.

EXAMPLE STEP 3:

The Ministry of Digital Technologies served as the policy lead, while IT Park managed implementation. Other actors included the Ministry of Economy (tax incentives), universities (innovation hubs), and regional governments (site provision and outreach). The evaluation revealed that local implementation varied widely – some hubs had dedicated managers and partnerships with local investors, while others were dormant due to unclear roles or underfunding. Additionally, youth and women-led businesses reported low awareness of support schemes outside major cities.

STEP 4: DESIGNING THE EVALUATION METHODOLOGY

Once the logic and actor landscape are clear, the evaluation team must select the most appropriate **evaluation criteria, approaches, and methods**. Unlike RIA, where cost-benefit analysis is often central, PIA tends to draw on a broader mix of qualitative and quantitative tools to assess results, relevance, and systemic performance.

Key decisions at this stage include:

- **Selecting criteria:** Most PIAs use the OECD DAC framework – relevance, effectiveness, efficiency, sustainability, impact, and coherence;
- **Defining evaluation questions:** These should reflect the policy's ambition, complexity, and level of maturity;
- **Choosing methods:** Theory-based evaluation, contribution analysis, outcome harvesting, and performance indicator tracking are commonly used;
- **Determining disaggregation:** Policies must be evaluated for how they affected different populations – particularly in terms of gender, geography, or income level.

EXAMPLE STEP 4: DESIGNING THE EVALUATION METHODOLOGY

The evaluation should combine three main tools:

- A survey of startup founders across six regions (disaggregated by gender and sector);
- Contribution analysis to examine the policy's role in startup success, relative to other market and institutional factors;
- Outcome harvesting to identify unintended but positive effects – such as new digital mentoring networks formed by alumni of the incubation centers.

Evaluation questions focused on policy reach (“Who accessed support?”), policy effectiveness (“Did startups grow, export, or survive at higher rates?”), and strategic coherence (“Was the policy aligned with national employment and education reforms?”).

STEP 5: DATA COLLECTION AND TRIANGULATION

Data collection in ex-post PIA is necessarily **multi-sourced and multi-layered**. Unlike regulations, where compliance or burden can be measured more discretely, policies often affect systems, populations, and institutions in ways that are difficult to capture through a single data stream. Evaluators must triangulate quantitative evidence, institutional data, and stakeholder perspectives to form a coherent picture of the policy’s performance.

Key sources include:

- **Administrative data** from ministries, agencies, or public platforms;
- **Monitoring data** (e.g., KPIs tracked during implementation);
- **Surveys and interviews** with beneficiaries, implementers, and policy designers;
- **Focus groups and case studies**, especially to understand unintended outcomes;
- **Independent or secondary data**, such as sector reports or national statistics.

Disaggregation is critical. Evaluators should ensure that data is broken down by gender, region, enterprise size, and other relevant dimensions – even if the policy itself did not originally include disaggregated tracking.

EXAMPLE STEP 5: DATA COLLECTION AND TRIANGULATION

To ensure robust and inclusive data collection, the evaluation team should combine administrative data from IT Park’s internal systems with complementary data sources such as firm-level export records from the Ministry of Economy. To assess accessibility and inclusion, the evaluation design should incorporate a structured survey targeting startup founders, with specific attention to capturing the experiences of women-led businesses, rural entrepreneurs, and other underrepresented groups. This should be complemented by site visits and semi-structured interviews at selected regional tech clusters, allowing for deeper insights into implementation dynamics and potential access disparities. By triangulating data from multiple sources – quantitative records, qualitative interviews, and survey responses – the evaluation will be better positioned to identify both achievements and blind spots, particularly in relation to geographic reach and equitable service provision.

STEP 6: ANALYSIS AND INTERPRETATION

In PIA, analysis must go beyond checking whether targets were met. It must explain **why certain outcomes occurred**, what factors contributed or hindered change, and how different groups experienced the policy. This requires returning to the reconstructed theory of change and testing it against the available evidence.

The evaluation team should:

- Apply DAC criteria systematically – but interpretively;
- Use **contribution analysis** to assess how policy elements interacted with broader dynamics;

- Identify **variability across regions, populations, or institutions**;
- Analyze coherence – did the policy align with other strategies or introduce fragmentation?
- Reflect on what worked, what didn't, and why.

STEP 6 EXAMPLE: ANALYSIS AND INTERPRETATION

The analysis revealed that startup formation had increased in target regions, particularly near universities that had active innovation hubs. However, retention and export performance varied. Startups that accessed both mentorship and tax benefits had significantly higher growth than those who only received space or basic training. Gaps in outreach and uneven application of eligibility criteria contributed to exclusion of rural and women-led firms. Some overlap was also identified between this policy and donor-funded programs with similar aims but fragmented tracking. These insights confirmed that the policy worked where conditions aligned, but highlighted serious gaps in capacity, coordination, and equity of access.

STEP 7: REPORTING, VALIDATION AND USE OF FINDINGS

A well-delivered PIA concludes not with a report, but with influence. The quality of findings matters – but their communication and uptake determine whether the evaluation has any lasting value. This means delivering results in ways that are technically credible, policy-relevant, and institutionally usable.

The final reporting package should include:

- A **concise summary report** with major findings and policy recommendations;
- **Annexes** with technical detail, raw data references, and methodological notes;
- A **visual dashboard or slide pack** for non-technical audiences;
- A **validation workshop** with government and stakeholder representatives;
- A **management response** and action plan, ideally endorsed at the leadership level.

STEP 7: REPORTING AND VALIDATION OF FINDINGS

Based on the PIA findings, the Ministry of Digital Technologies may choose to convene a multi-stakeholder roundtable with regional innovation hub coordinators, private accelerators, and development partners. Key findings could be presented through visual dashboards, highlighting policy uptake by region, demographics of supported founders, and identified performance gaps. Informed by the evaluation, a set of strategic responses may be proposed, such as:

- *Launching a national mentorship platform with targeted support for underserved regions;*
- *Revising eligibility criteria to enhance clarity, accessibility, and equity;*
- *Establishing a cross-sector monitoring working group to improve data consistency and institutional learning.*

By integrating evaluation results into decision-making processes, the PIA can serve as a vehicle for continuous improvement – supporting more inclusive and effective digital entrepreneurship policies in future programming cycles.

8. CONCLUDING REMARKS

This **Methodology Guide** has been developed to support the institutionalization of robust, credible, and adaptive M&E practices in Uzbekistan, with a particular focus on strengthening policy and regulatory governance related to the country's digital economy transformation. It complements the forthcoming M&E Decree by providing **practical guidance** on how ministries, agencies, and oversight bodies can operationalize monitoring and evaluation systems that contribute directly to the implementation of Digital Uzbekistan 2030 and related strategic initiatives.

Building an effective M&E ecosystem for the digital economy is not the work of a single reform or decree. It requires **sustained leadership, improved technical capacities, and a shift towards a governance model where evidence routinely informs policy**, regulation, program design, and resource allocation in digital development. This Guide provides a roadmap for this process outlining foundational concepts of results-based management, the necessary preconditions for effective monitoring and evaluation, the planning responsibilities of the center of government, and technical methodologies that practitioners can apply across digital economy initiatives.

By using the tools, templates, and methodologies presented here, institutions will be better positioned to:

- **Develop coherent monitoring and evaluation frameworks** specifically geared to the needs of digital transformation policies and programs;
- **Conduct rigorous ex-post evaluations** of regulations and policies that support ICT entrepreneurship, digital skills development, innovation infrastructure, and e-governance;
- **Build integrated monitoring systems** that track not only activity delivery but the achievement of outcomes such as digital inclusion, startup growth, and innovation competitiveness;
- Ensure that **monitoring and evaluation outputs are used** to guide strategic adjustments in digital economy planning and reform.

