IFAD

Investing in rural people

IFAD REVISED EVALUATION MANUAL

PART I

Interim version

IFAD

Investing in rural people

IFAD REVISED EVALUATION MANUAL

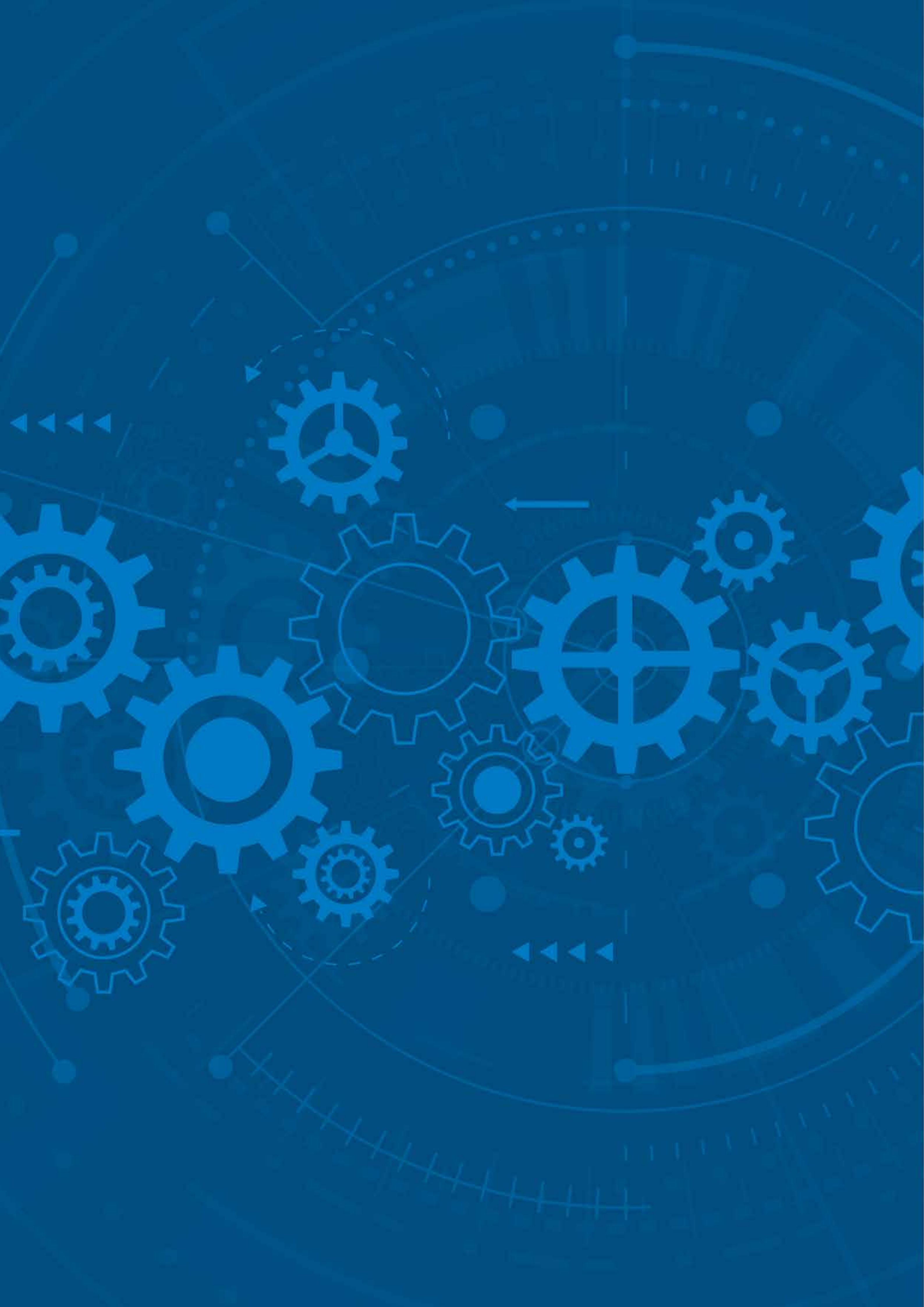
PART I

Interim version

ISBN 978-92-9266-265-3

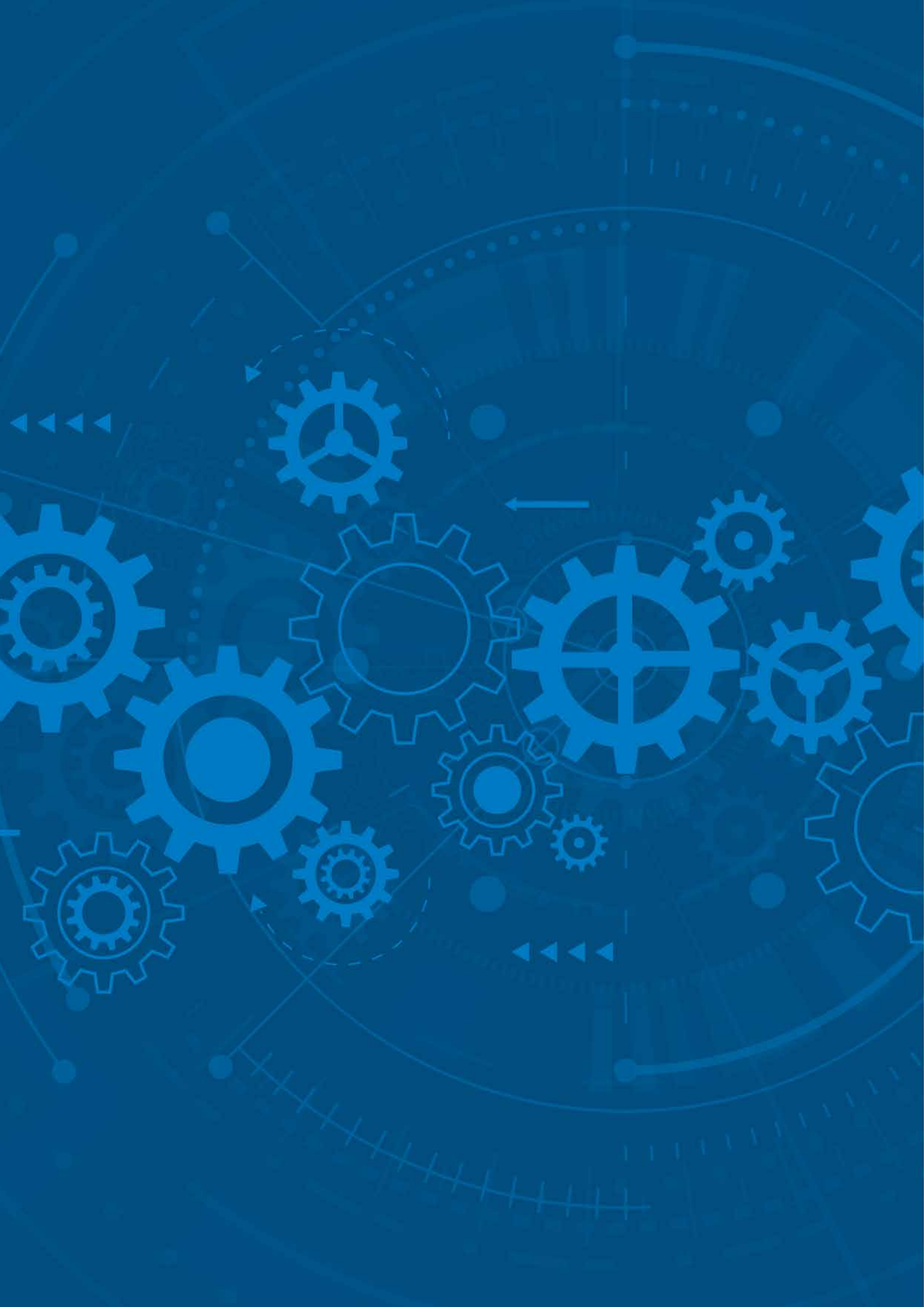
May 2022
Graphic designer: Laura Monopoli

Contents



Abbreviations and acronyms

CI	core indicator
COI	core outcome indicator
COSOP	country strategic opportunities programme
DEF	Development Effectiveness Framework
ECG	Evaluation Cooperation Group
GIZ	German Agency for International Cooperation
ICT	information and communications technology
IEG	Independent Evaluation Group
IOE	Independent Office of Evaluation of IFAD
M&E	monitoring and evaluation
OECD-DAC	Organisation for Economic Co-operation Development – Development Assistance Committee
OPR	Operational Policy and Results Division
PCR	project completion report
PMU	project management unit
QCA	qualitative comparative analysis
RIA	Research and Impact Assessment Division
RMF	Results Management Framework
SDG	Sustainable Development Goal
SKD	Strategy and Knowledge Department
ToC	theory of change
UNEG	United Nations Evaluation Group



Foreword

The 2022 edition of the IFAD Evaluation Manual was prepared in consultation with IFAD Management at the recommendation of the 2019 External Peer Review of IFAD's Evaluation Function. This is the third revision of the manual. The first edition was released in 2009 and the second in 2015.

This revised manual follows recognized international practices and standards, notably those of the United Nations Evaluation Group, the Evaluation Cooperation Group of the Multilateral Development Banks and the Evaluation Network of the Organisation for Economic Co-operation Development. We are deeply grateful to representatives from these networks who have collaborated with us to prepare the manual.

The manual draws on contemporary literature on evaluation and advances in evaluation practices since the launch of the 2030 Agenda for Sustainable Development. Examples of evaluation approaches and perspectives linked to the 2030 Agenda are:

i. The notion of **transformative change**. Evaluators are increasingly called on to assess effects on norms and systems beyond the immediate project results.

ii. The perspective of **complexity and systems thinking**. The solutions to poverty, inequality, and other global challenges are deeply intertwined. Understanding these interactions requires the use of systems analysis.

iii. Addressing **sustainability and climate resilience**. The Sustainable Development Goals identify climate change as a multiplier of threats that are capable of hindering progress on poverty, hunger, equality and health. Evaluators need to assess the ability of countries, communities and households to cope with unpredictable climate shocks.

iv. **Social justice and gender intersectionality**. The commitment to leave no one behind is at the heart of the 2030 Agenda. Evaluations need to consider the intersection of gender and other socio-economic characteristics, such as ethnicity, age, caste and income level to address multiple causes of discrimination and exclusion.

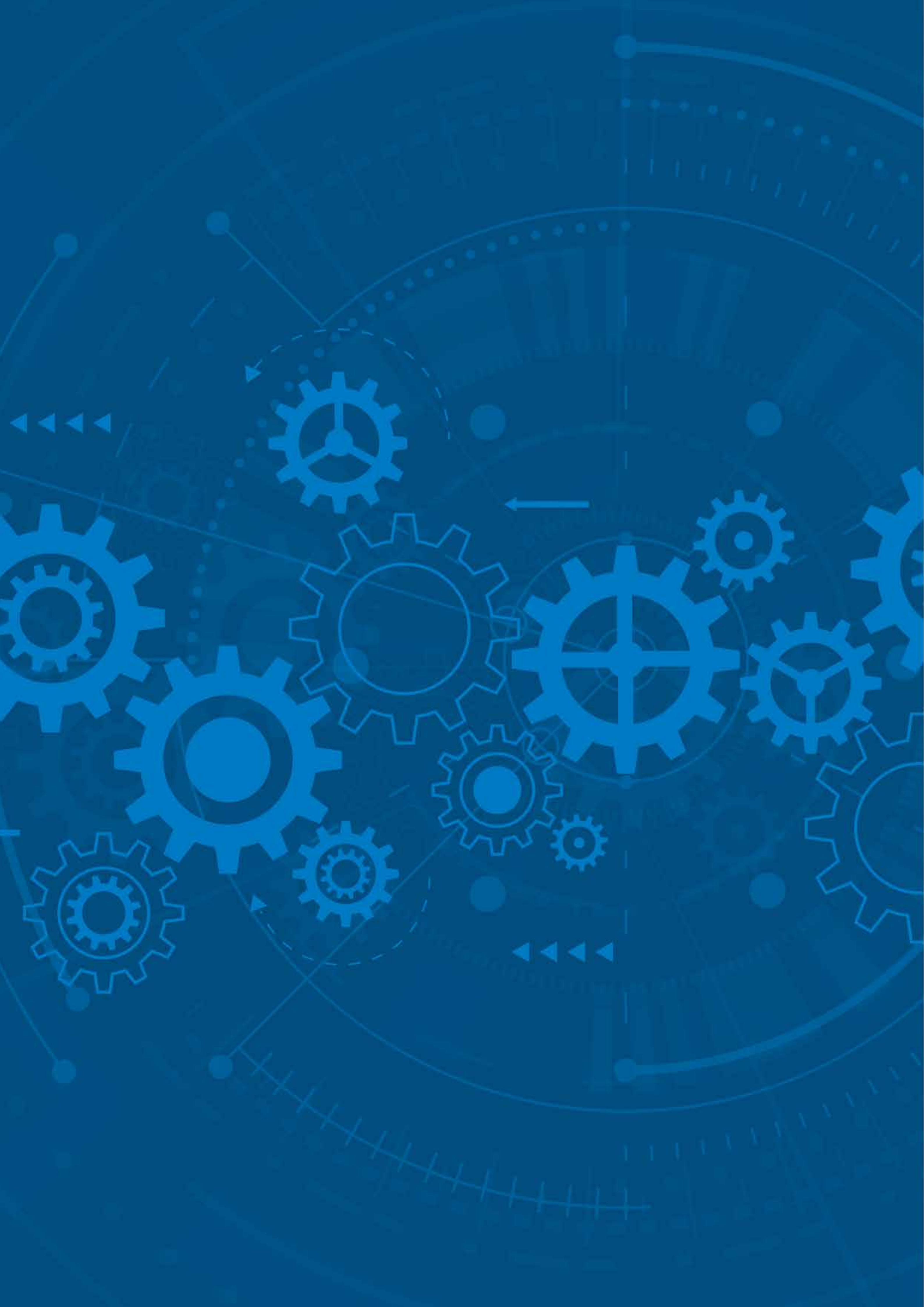
For the first time, the IFAD Evaluation Manual applies to both self- and independent evaluation and provides methodological guidance and standards for evaluations across the organization. Our intention is to establish common quality standards for self- and independent evaluation, enhance the consistency between the two and, ultimately, foster a stronger results and evaluation culture at IFAD.

The purpose of this manual is not to provide a set of rigid and immutable directives. It is a living document that draws from the wealth of experience and good practices at IFAD in evaluating rural development and poverty reduction programmes as well as corporate policies and processes. It is issued in electronic form, which will facilitate periodic revisions and updates.

While users of this manual will typically include IFAD staff members and consultants, we trust that it will be of interest to a broader group of stakeholders working for governments, the private sector, civil society and other development partners who are committed to rural development and poverty reduction. A companion online training course and dedicated learning events are being prepared to help familiarize potential users with the manual.

Indran A. Naidoo, PhD
Director
Independent Office of Evaluation

Donal Brown
Associate Vice-President
Programme Management Department



Acknowledgements

This manual was prepared by a task force comprising staff from the Independent Office of Evaluation of IFAD (IOE), Operational Policy and Results Division (OPR) and the Research and Impact Assessment Division (RIA), under the general guidance of: Indran Naidoo, Director, IOE; Donal Brown, Associate Vice-President, Programme Management Department (PMD); Meike van Ginneken, former Associate Vice-President, Strategy and Knowledge Department (SKD); and, more recently, Jyotsna Puri, Associate Vice-President, SKD. The task force was led by Fabrizio Felloni, Deputy Director, IOE; Nigel Brett, Director, OPR; Sara Savastano, Director, RIA. Thomas Eriksson, former Director, OPR, co-led the task force in its initial stages. Other members of the task force included: Simona Somma, Evaluation Officer, IOE, who oversaw the management of the process; Lauren Phillips, former Lead Policy Adviser, OPR; Aslihan Arslan, Senior Economist, RIA; Romina Cavatassi, Lead Economist, RIA; Dimitra Stamatopoulos, Policy and Results Specialist, OPR; Xiaozhe Zhang, Policy and Result Specialist, OPR; Rodney Cooke, consultant, OPR; and Emanuele Zucchini, consultant, RIA.

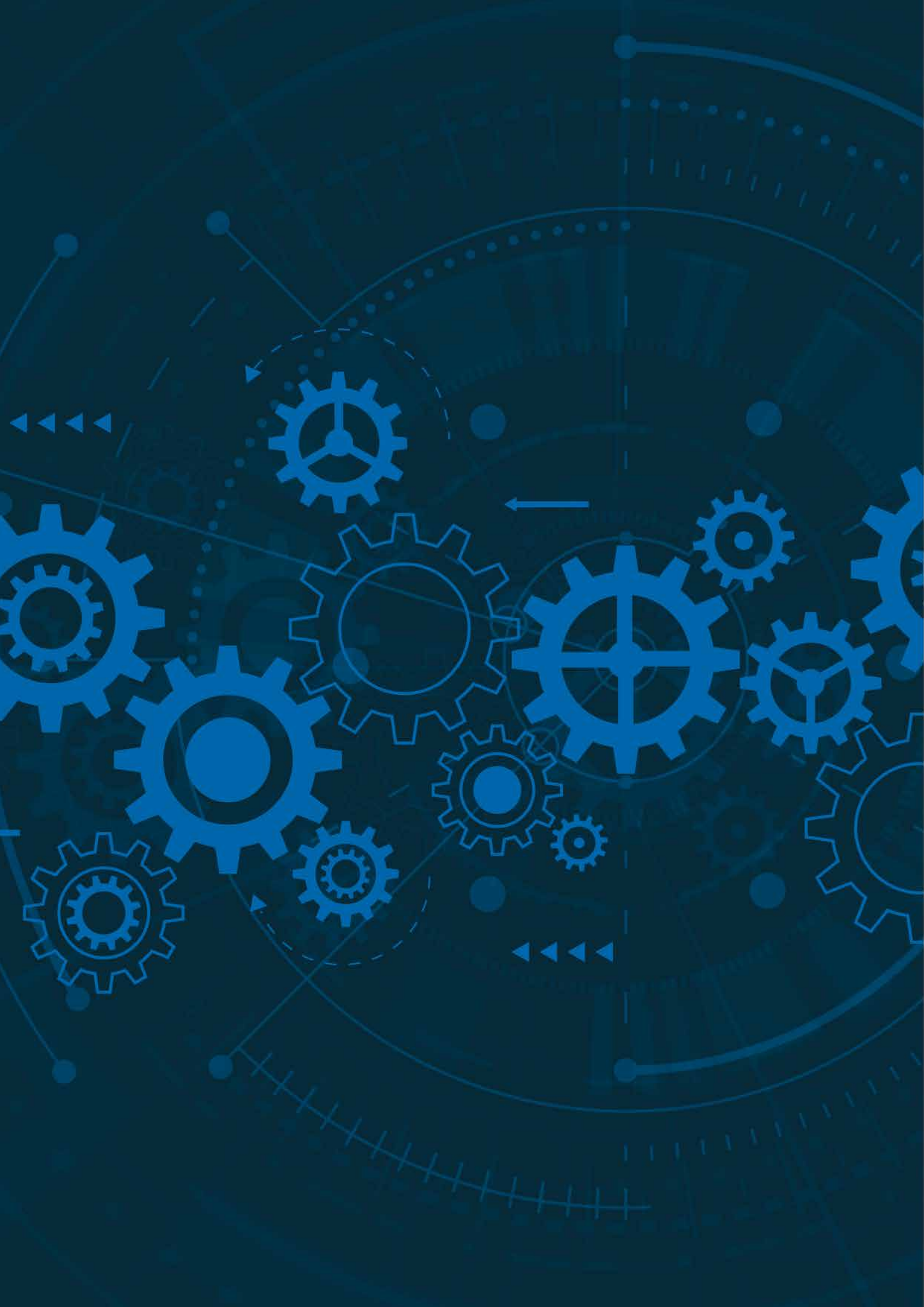
Paula Silva Villanueva, senior consultant, provided extensive technical advice in drafting part 1 of the manual. Dorothy Lucks, senior consultant, provided comments on part 1 and advice to IOE in preparing relevant chapters of part 2.

The report benefited from substantive review by the Evaluation Advisory Panel of IOE, comprising: Bagele Chilisa, University of Botswana; Gonzalo Hernández Licona, University of Oxford; Hans Lundgren, former Head Evaluation Secretariat, Organisation for Economic Co-operation and Development – Development Assistance Committee (OECD-DAC) Evaluation Network; Donna Mertens, Professor Emeritus, Gallaudet University; and Rob van den Berg, Visiting Professor, King's College London and former Director of the Independent Evaluation Office, Global Environment Facility.

In June 2021, a draft version of the manual was discussed with an international group of experts co-appointed by IOE and Management. It comprised: Marie Gaarder, Executive Director, International Initiative for Impact Evaluation; Stefano Gagliarducci, Professor, Department of Economics and Finance, Tor Vergata University of Rome; Megan Kennedy-Chouane, Head of the Evaluation Unit, OECD-DAC Development Co-operation Directorate; Sebastian Stolorz, Senior Operations Officer, Strategy, Risk and Learning Department, the World Bank; Maya Vijayaraghavan, Methods Advisor, Independent Evaluation Department, Asian Development Bank. In addition, Andrea Cook, Director of Evaluation, World Food Programme (WFP), provided written comments.

Contributions from many staff members in IOE, PMD and SKD are also acknowledged.

Introduction



Introduction

A. Why this manual?

1. IFAD is committed to making a significant contribution to eradicating poverty and hunger in rural areas of developing countries while positively impacting gender equality, climate and social justice. The main purpose of the IFAD Revised Evaluation Manual (2021) is to ensure the quality, consistency, rigour and transparency of the evaluation function at IFAD to increase the effectiveness of IFAD efforts to contribute to the well-being of poor people in rural areas.

2. This manual is a living electronic document that will be adapted over time to reflect evolving practice, needs and lessons. The IFAD Revised Evaluation Manual represents a major revision of the 2015 edition. The changes in the latest version are designed to improve the implementation of IFAD's evaluation policy, with which it is aligned. The manual seeks to revamp, update and consolidate current guidelines. For the first time, it provides a comprehensive institution-wide approach through which self- and independent evaluations are planned, conducted and used. The manual places renewed emphasis on the importance of harmonizing and streamlining the two and maximizing the use of findings and lessons when planning and implementing projects and programmes.

3. The revision of the 2015 IFAD Evaluation Manual was undertaken by the Independent Office of Evaluation of IFAD (IOE) and Management in recognition of the dynamic environment in which IFAD operates and in response to evolving approaches and methodologies of international development evaluation practices. It will help to ensure that IFAD's methodological practice remains state-of-the-art.

B. For whom is this manual written?

4. The evaluation manual sets standards for self- and independent evaluations at IFAD. Its main audience is IFAD's staff and consultants who manage evaluations or are involved in independent- and self-evaluation processes.

5. The manual's secondary audience includes IFAD's stakeholders and partners, such as governments, the private sector, civil society and development partners involved in IFAD's planning, monitoring and evaluation processes.

6. The manual may also be of interest to individuals outside IFAD involved in the evaluation of rural development programmes. This includes Member States, international organizations, national non-governmental partners and beneficiaries and rural development practitioners.

C. What does it contain?

7. The manual describes how evaluation is performed at IFAD and is therefore not a sourcebook on evaluation in general. It includes basic guidance on evaluation fundamentals and the criteria applicable to all evaluations. It also contains a section on organizational learning, recognizing that reports are of limited value if the knowledge they contain is not appropriately used by as many people as possible. To this end, it provides details on the complementarities of IFAD's self- and independent-evaluation systems and related evaluation products with a view to strengthening accountability, learning and the overall utility of evaluations efforts. It also comprises specific methodological guidance on all evaluation products.

8. Readers are encouraged to read all chapters to develop a good understanding of how evaluations should be interpreted, managed, conducted and used. The manual also serves as a reference for information about specific issues and evaluation products.

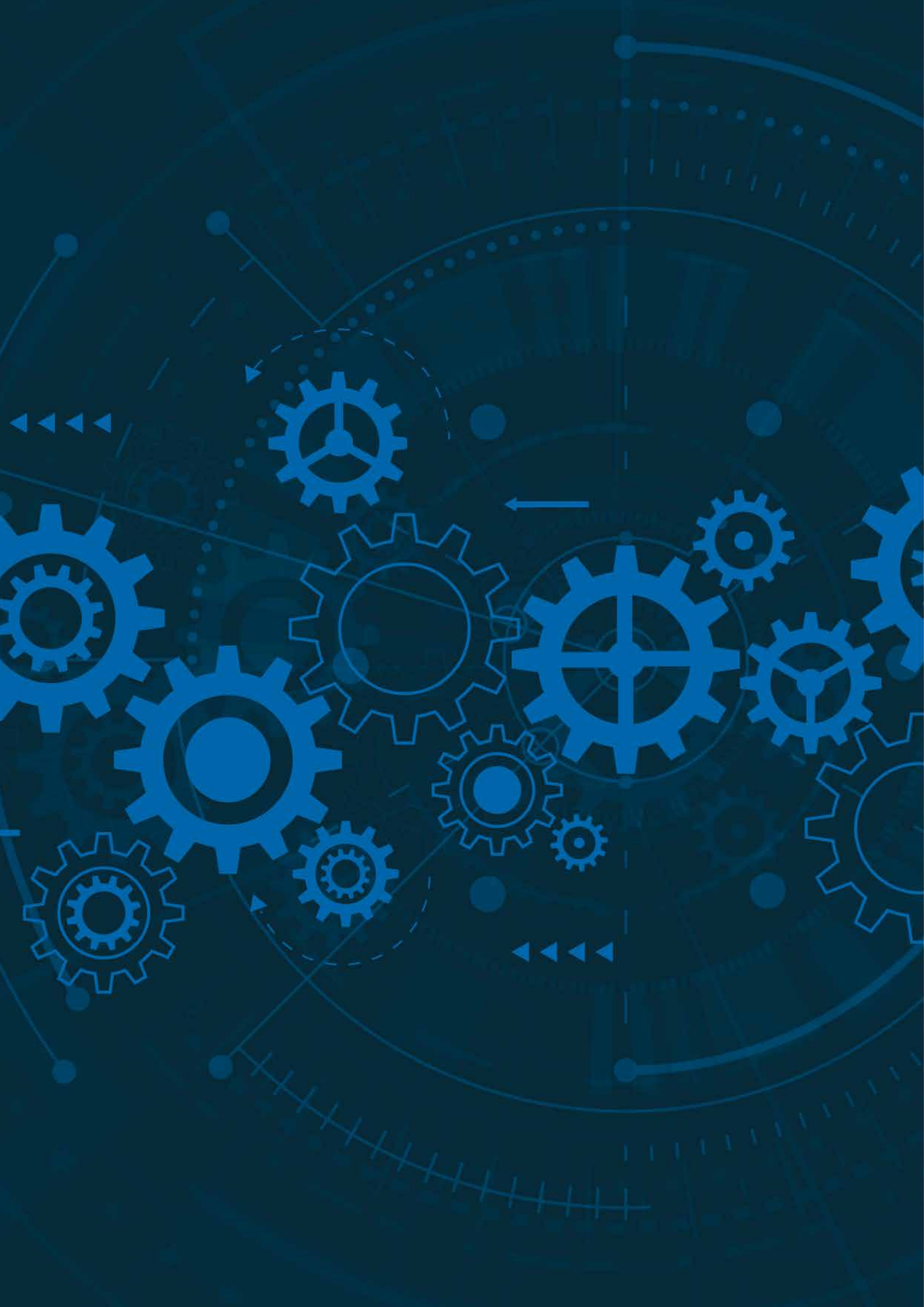
9. The manual is divided into two parts:

10. Part 1 (chapters I-III) provides the overall context for evaluation in IFAD. It covers a number of foundational elements, including IFAD's mission, evaluation objectives, architecture, frameworks and principles and criteria that guide all evaluations within the Fund.

11. Part 2 (chapters IV and onwards) provides practical guidance on various self- and independent-evaluation products. These individual chapters can be used in sequence or as individual sections and are intended to be living, continuously evolving documents to support evaluation in IFAD.

Part 1

Evaluation in IFAD

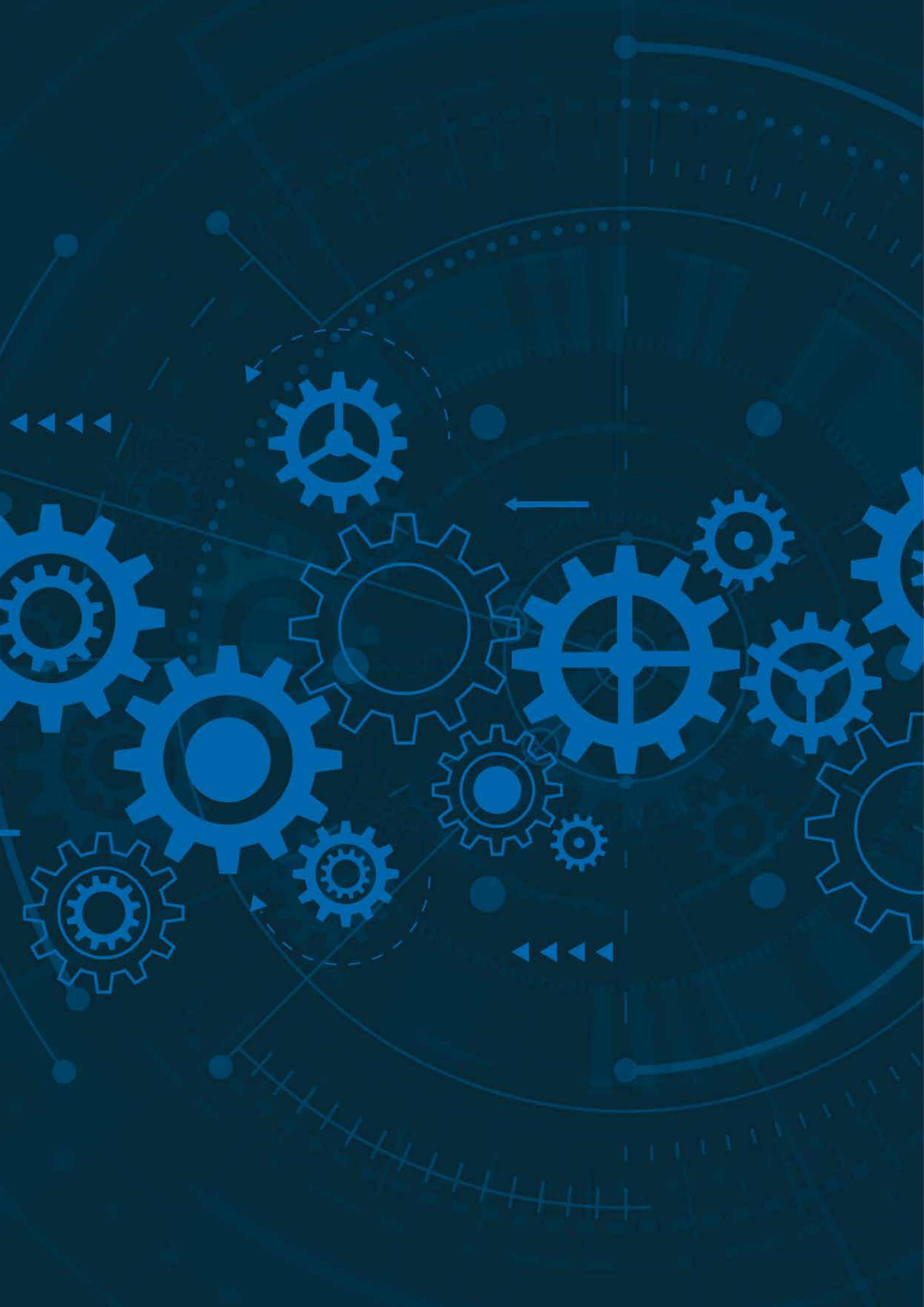


Part 1. Evaluation in IFAD

1. Part 1 introduces the foundations for evaluation in IFAD and comprises three chapters. Chapter I puts IFAD's evaluation efforts in the context of IFAD's endeavour to contribute to inclusive and sustainable rural transformation and to the achievement of the United Nations 2030 Agenda for Sustainable Development. It presents IFAD's theory of change (ToC) and introduces the purpose of evaluation and its overarching principles. Chapter II presents IFAD's institution-wide evaluation and learning systems, different functions and types of evaluation. Chapter III explains the methodological fundamentals applicable to all evaluations.

Chapter I

Laying the foundations



I. Laying the foundations

2. The 2030 Agenda unequivocally amplifies the call for greater attention, cooperation and investment in rural development (box 1). IFAD contributes to lifting poor rural people out of poverty. No poverty eradication and inclusive growth agenda can succeed without serious attention to rural areas and sectors, which support the livelihoods of small-scale producers. Indeed, poverty has multiple dimensions that go beyond low levels of income, consumption and material assets; this is why IFAD targets its investments toward inclusive rural transformation, which is a comprehensive process during the structural transformation of economies that has social as well as economic implications.[1]

1 Rural Development Report 2016: Fostering inclusive rural transformation.

BOX 1

Excerpts from the preamble of the 2030 Agenda

We are resolved to free the human race from the tyranny of poverty and want and to heal and secure our planet. We are determined to take the bold and transformative steps which are urgently needed to shift the world onto a sustainable and resilient path. As we embark on this collective journey, we pledge that no one will be left behind. The 17 Sustainable Development Goals (SDGs) and 169 targets which we are announcing today demonstrate the scale and ambition of this new universal Agenda. They seek to build on the Millennium Development Goals and complete what these did not achieve. They seek to realize the human rights of all and to achieve gender equality and the empowerment of all women and girls. They are integrated and indivisible and balance the three dimensions of sustainable development: the economic, social and environmental. […]

We will devote resources to developing rural areas and sustainable agriculture and fisheries, supporting smallholder farmers, especially women farmers, herders and fishers in developing countries, particularly least developed countries.

Source: https://www.un.org/development/desa/dspd/2030agenda-sdgs.html

A. The role of IFAD

3. The mission of IFAD is to facilitate both public and private investment, support national and global policy making, generate and share knowledge and develop partnerships, all in pursuit of transforming agriculture, rural economies and food systems to make them more inclusive, productive, resilient and sustainable.

4. IFAD is a specialized agency of the United Nations and an international financial institution (IFI) focused exclusively on reducing poverty and food insecurity in rural areas through agricultural and rural development. IFAD's portfolio targets small-scale producers, owners of small- and medium-sized rural enterprises and vulnerable rural groups, such as women, youth, indigenous peoples and persons with disabilities.[2]

5. IFAD has been ranked the top development cooperation agency (among 49 institutions) by the Center for Global Development in its QuODA 2021 (Quality of Official Development Assistance), based on an assessment of four dimensions of quality: prioritization; ownership; transparency and untying; and evaluation. The ranking specific to the evaluation dimension placed IFAD as the fourth development cooperation agency in terms of the quality of providers' learning and evaluation systems.[3]

6. Through its Strategic Framework 2016-2025[4] IFAD is committed to pursuing three interlinked strategic objectives: (i) increase poor rural people's productive capacities; (ii) increase poor rural people's benefits from market participation; and (iii) strengthen the environmental sustainability and climate resilience of poor rural people's economic activities. IFAD12 (2022-2024) is the last full replenishment cycle[5] that will operate under the current strategic framework.[6]

7. The Twelfth Replenishment of IFAD's Resources[7] (IFAD12) is a critical cycle for IFAD to increase its contributions to SDGs and deliver on its core mandate of promoting sustainable rural transformation.[8] IFAD's vision is built upon a ToC that is articulated in the context of IFAD12 and included in IFAD's Results Measurement Framework (RMF12).[9]

8. The pathways to impact are represented in figure 1. The ToC is a conceptual model, not a literal representation of a linear process. Its main purpose is to provide a conceptual framework for understanding important changes that IFAD must make for long-term success.

2 The main instruments for delivery are loan-funded projects and programmes, which IFAD helps governments, beneficiaries and other stakeholders develop and implement. IFAD also has a small grant-financing window, and a new grant policy has just been developed, along with reimbursable technical assistance. Moreover, IFAD is the first fund in the United Nations system to receive a public credit rating (AA+ by Fitch Ratings and Standard & Poor's), which will allow IFAD to strengthen its resource base and catalyse private sector finance toward achievement of the SDGs.

3 QuODA 2021: Aid Effectiveness Isn't Dead Yet: **https://www.cgdev.org/publication/quality-official-development-assistance**

4 **IFAD Strategic Framework 2016-2025.**

5 IFAD's core financing is drawn from several sources, including contributions from Member States and other donors, investment income and loan reflows every three years. Based on these financial resources, IFAD operations are planned for a three-year replenishment period. IFAD12 is the twelfth replenishment cycle, covering the period 2022-2024 (for more details about IFAD12 see: **https://www.ifad.org/en/ifad12/)**

6 The SDGs and IFAD Strategic Framework 2016-2025 are the reference documents that indicate IFAD's longer-term ambitions.

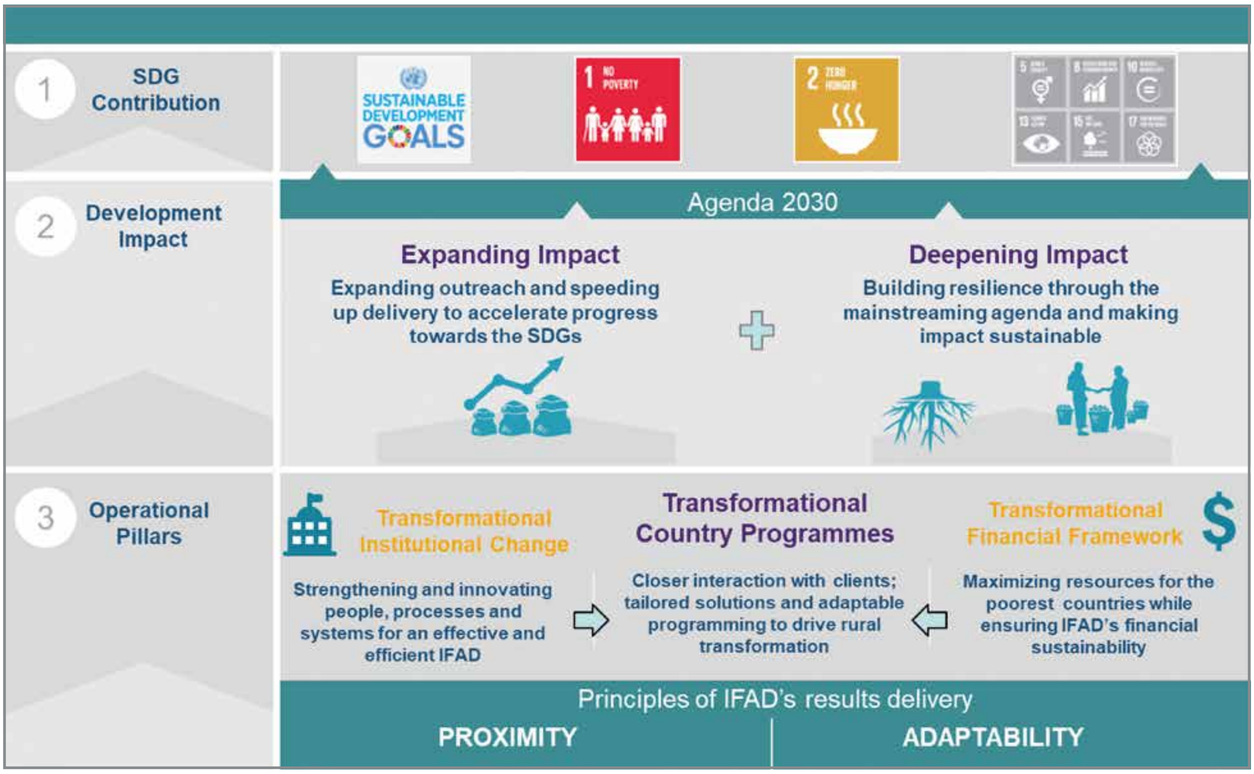
7 **https://www.ifad.org/en/ifad12/**

8 For IFAD12, IFAD set a target of reaching 127 million people with its operations. See Report of the Consultation on the Twelfth Replenishment of IFAD's Resources – Recovery, Rebuilding, Resilience. December 2020. This is the first time that IFAD will report against a specific ToC. The ToC and the relevant sections in the evaluation manual will be revised when the new replenishment begins.

9 This is the first time that IFAD is reporting against a specific ToC. The ToC and relevant sections in the manual will be revised when the new replenishment begins.

FIGURE 1

IFAD12 theory of change



Source: IFAD (2021). Report of the Consultation on the Twelfth Replenishment of IFAD's Resources.

9. Tier 1 – SDG contribution. IFAD maintains its ambition to make significant contributions to SDG 1 (no poverty) and SDG 2 (zero hunger), tracked by measures of extreme poverty, food insecurity and the productivity of small-scale producers, while positively impacting the broader development goals, especially those focused on gender equality, climate and social justice.

10. Tier 2 – Developmental impact. Here, it is assumed that success in contributing to global poverty reduction and food security targets is achieved by country-level outcomes and impact. To be successful, IFAD must:

- **Expand impact**: Increase outreach and speed up delivery to accelerate progress toward ending poverty and hunger.
- **Deepen impact**: Target the poorest and most vulnerable groups, strengthen systems and people's resilience to shocks and stressors and ensure that impact is sustainable. Environment and climate change, gender, nutrition and youth are critical and intersecting areas of the work toward reducing poverty and hunger and fostering resilient rural livelihoods. Thus, IFAD identifies four mainstreaming themes – youth, gender, nutrition and climate – as central to deepening impact and transforming the lives of rural populations. In practice, the mainstreaming agenda means that the ToC of projects at the design phase needs to clearly show synergies and intersectionality between different mainstreaming themes.

11. IFAD's developmental impact relates to the impact of IFAD-funded operations and is measured through independent evaluations,[10] Management's impact assessments and monitoring of outputs and outcomes within IFAD's core indicators framework. Yet, the Fund is situated in a wider global development policy context that is complex, contested and non-linear. Evaluation efforts must take into account the web of actors, conflicting interests and systems interactions that enable or constrain IFAD's impact (implications for evaluation are presented in chapter III).

10 IFAD is a results-driven organization, as evidenced by the focus on results introduced in its first Development Effectiveness Framework (DEF). As the institution diversified its instruments and heightened its ambitions, it also updated its DEF in 2021. The updated DEF will capture evolving priorities and new areas of work to ensure that the institution's approach to results is all-encompassing. IFAD's success will be assessed against the agreed indicators of the IFAD12 RMF.

12. Tier 3 – Operational pillars. The IFAD12 ToC highlights the fact that transformational country programmes are needed to drive transformative results for poor rural people. To achieve meaningful impact, country programmes must: (i) include inclusive approaches aimed at leaving no one behind; (ii) deepen partnerships and enhance government ownership through a suite of adaptive products and tools suited to country needs; and (iii) make significant investment in innovation and risk.

13. **The focus on resilience to shocks and stressors, as well as transformational change, requires evaluators to have an operational definition of transformational results**. This is not explicitly defined in IFAD12 but can be derived from it[11] and complemented with definitions provided by the Independent Evaluation Group (IEG) of the World Bank[12] and the German Agency for International Cooperation (GIZ).[13] These sources are established references in international evaluation; combine social, environmental and climate change resilience aspects; and place the importance of systemic changes at the forefront.

14. **Transformational results** are those that lead to a deep, systemic and sustainable change with the potential for large-scale impact at the national or global level. Transformative results ultimately generate changes that are profound enough to shift societies onto fundamentally different development pathways, turning a current (ecological, social, political, economic, scientific or technological) system into a fundamentally new one that forms the new mainstream.

15. **Proximity and adaptability** underpin the ToC and are cross-cutting principles for delivering results. Focusing on these two principles will enable IFAD to expand and deepen the results achieved when working through country programmes.

- **Proximity**: Working closer to all its partners and members of marginalized and vulnerable communities to facilitate the ability to work in genuine partnership, find solutions to common problems and make a more transformational impact on policy.

- **Adaptability**: Taking an adaptive approach to "doing development." Adaptive management approaches emphasize the ability to proactively and reactively learn, respond to changes and evolve quickly. Rather than adopting fixed targets as a results-based approach would, project teams and governments should be encouraged to adapt the means to achieving end goals based on lessons learned along the way.

16. **Consequently, evaluation efforts[14] must support IFAD in making course corrections more quickly when risks emerge that could undermine development objectives and outcomes, or when economic or other shocks occur**.[15] Both principles call for more frequent planning with stakeholders and greater agility during project design and implementation. This reinforces the role of evaluation, not only to assess IFAD's impact and the role of partners but to navigate uncertainty and the complexity of operations, understand the pathways to impact and examine and assess what works, for whom, where, when and why. Chapter III presents IFAD's evaluation approach.

11 According to IFAD12, conditions for achieving transformational results are: (i) focus on people's resilience to ensure sustainability and impact even in a crisis; (ii) nurturing partnership with governments, the private sector, civil society and the non-governmental organization community, think tanks and other development organizations. The Rural Development Report 2016 also provides a definition for rural and structural transformation.
https://www.ifad.org/en/web/knowledge/-/rural-development-report-2016-fostering-inclusive-rural-transformation

12 IEG World Bank (2016) Supporting Transformational Change for Poverty Reduction and Shared Prosperity.

13 GIZ (2019) Transformative Project Design.

14 IFAD monitoring system plays a fundamental role in support of the evaluation function. This manual focuses on evaluation efforts.

15 IFAD12: Business Model and Financial Framework 2022-2024.

B. Evaluation's role in supporting IFAD in achieving its development objectives

17. Evaluation plays an important role in IFAD's business model, including operational focus, corporate processes, accountability and learning systems. They are refined, adjusted and sharpened through feedback from evaluation to ensure that the Fund is in the best possible position to fulfil its mandate and meet its corporate goals.

18. The 2021 Revised IFAD Evaluation Policy constitutes the overall framework for evaluations within the institution.[16] It outlines the roles and responsibilities in evaluation and includes IFAD's evaluation ToC. Chapter II presents and develops the key tenets of IFAD's evaluation and learning system, including its various functions, types of evaluation and feedback loops.

19. In supporting IFAD in achieving greater development effectiveness, the evaluation policy identifies two primary purposes:

- Promote accountability by providing an evidence-based assessment of the outcomes achieved through IFAD lending and non-lending support and by putting in place the necessary corporate business model, policies, strategies and guidance, as well as resources and capacities to achieve these results; and
- Contribute to enhanced learning, knowledge management and transparent feedback mechanisms to improve current and future policies, strategies, programmes, projects and processes (figure 2).

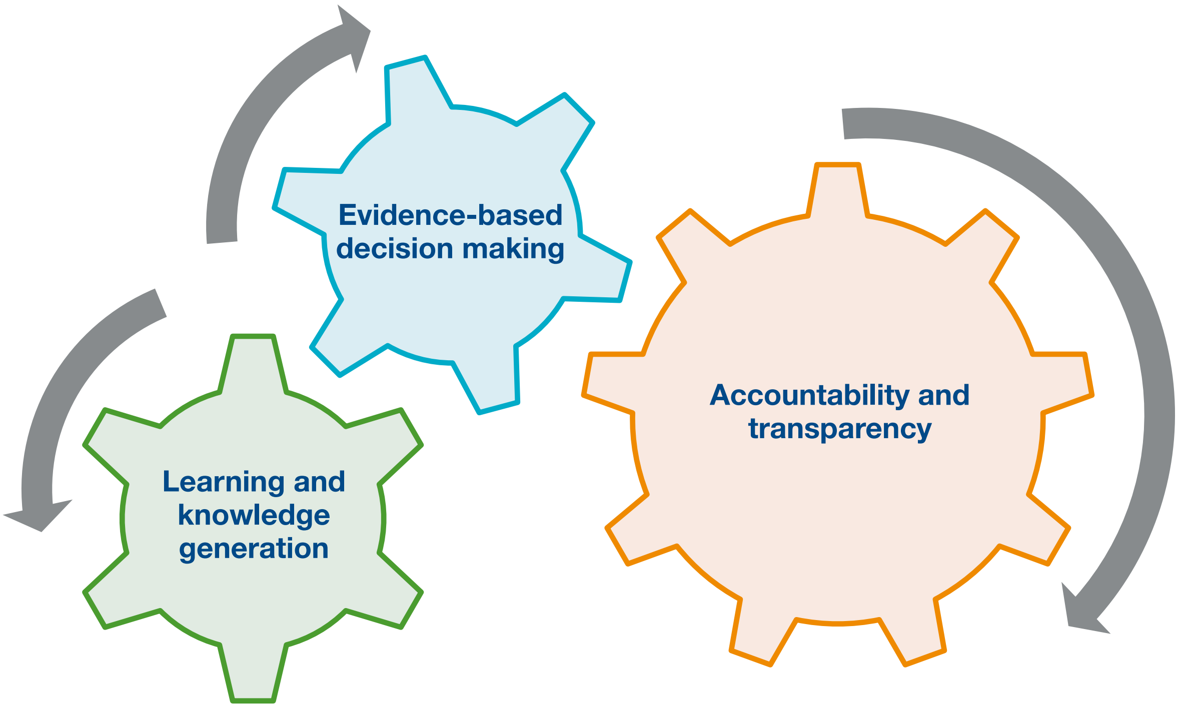
20. The 2019 report of the external peer review panel on the evaluation function at IFAD emphasized that while accountability was a strong point of IFAD's evaluation function, learning could be improved, as could the incentives and methods through which learning loops were deployed across the institution.[17] Nonetheless, evaluation helps IFAD improve the planning and implementation of its policies, strategies and operations by determining the relevance and fulfilment of its development objectives. As illustrated in figure 2, both accountability and learning functions aim to promote a results-based culture, evidence-based planning and adaptive management across IFAD.

16 Revised IFAD Evaluation Policy, 2021. https://webapps.ifad.org/members/eb/132/docs/EB-2021-132-R-5-Rev-1.pdf

17 https://webapps.ifad.org/members/eb/124/docs/EB-2018-124-R-8.pdf

FIGURE 2

Accountability and learning for evidence-based decision-making



21. As further explained in chapter II, IFAD's evaluation architecture includes the independent and self-evaluation systems. IOE conducts independent evaluations, whereas IFAD's self-evaluation system is the responsibility of Management. The two systems work jointly through the harmonization of processes and consultations at key stages of evaluations, consistent with the independence of IOE. The range of IFAD evaluation products (see part 2 of this manual) ensures that both dimensions of evaluation are adequately covered.

C. Core evaluation principles

22. The 2021 Revised IFAD Evaluation Policy identifies six interrelated key principles that underpin the organization's approach to evaluation and provide the conceptual framework for evaluations. These principles are: usefulness; impartiality and credibility; transparency, partnership, consultation and collaboration; evaluability; and value for money/cost-effectiveness.[18]

23. IFAD also subscribes to the overarching norms and standards adopted by the United Nations Evaluation Group (UNEG),[19] the Evaluation Cooperation Group (ECG) of the Multilateral Development Banks[20] and the quality standards and principles of the Organisation for Economic Co-operation Development (OECD).[21] Consistency and compliance with these broader principles are at the core of IFAD's evaluative work.

18 https://webapps.ifad.org/members/eb/132/docs/EB-2021-132-R-5-Rev-1.pdf

19 UNEG Norms and Standards for Evaluation, 2016. http://www.unevaluation.org/document/detail/1914

20 Evaluation Cooperation Group's Good Practice Standards for the Evaluation of Public Sector Operations, 2012. https://ecgnet.org/content/public-sector-operations

21 OECD-DAC. https://www.oecd.org/dac/evaluation/
Quality standards: https://www.oecd.org/development/evaluation/qualitystandards.pdf
Principles: https://www.oecd.org/development/evaluation/2755284.pdf

24. Evaluation ethics is the fundamental principle underlying the six mentioned above in evaluation practice. It is the responsibility of the evaluation team to uphold ethical codes of practice, guidelines and principles. UNEG defines ethics as "the right or agreed principles and values that govern the behaviour of an individual within the specific, culturally defined context within which an evaluation is commissioned or undertaken," and identifies four key principles (see figure 3): integrity; accountability; respect; and beneficence. Systematic attention to these principles helps balance the goals of evaluations and those who conduct them with the rights and interests of diverse participants and their communities. In contrast, failure to systematically consider ethics throughout the evaluation cycle can have adverse consequences for intended beneficiaries of the evaluation.

FIGURE 3

IFAD evaluation function endorses UNEG's ethical principles for evaluation

INTEGRITY

the active adherence to moral values and professional standards, which are essential for responsible evaluation practice.

- Honesty and truthfulness in communication and actions.
- Professionalism based on competence, commitment, ongoing reflective practice and credible and trustworthy behaviour.
- Independence, impartiality and incorruptibility to mitigate or prevent conflicts of interest, bias or undue influence of others, which may otherwise compromise responsible and professional evaluation practice.

ACCOUNTABILITY

the obligation to be answerable for all decisions made and actions taken; to be responsible for honouring commitments, without qualification or exception; and to report potential or actual harms observed through the appropriate channels.

- Transparency of the evaluation, thereby increasing accountability for performance to the public.
- Responsiveness as questions or events arise. Where corruption, fraud, sexual exploitation or abuse or other misconduct or waste of resources is identified, it must be referred to appropriate channels.
- Taking responsibility for meeting the evaluation purpose and for actions taken.
- Justifying and fairly and accurately reporting decisions, actions and intentions to stakeholders.

RESPECT

involves engaging with all stakeholders of an evaluation in a way that honours their dignity, well-being and personal agency while being responsive to their sex, gender, race, language, country of origin, LGBTQ status, age, background, religion, ethnicity and ability and to cultural, economic and physical environments.

- Access to the evaluation process and products by all relevant stakeholders – with due attention to factors that can impede access such as sex, gender, race, language, country of origin, LGBTQ status, age, background, religion, ethnicity and ability.
- Meaningful engagement and fair treatment of all relevant stakeholders in the evaluation processes, so they can actively inform the evaluation approach and products rather than being solely a subject of data collection.
- Fair representation of different voices and perspectives in evaluation products.

BENEFICENCE

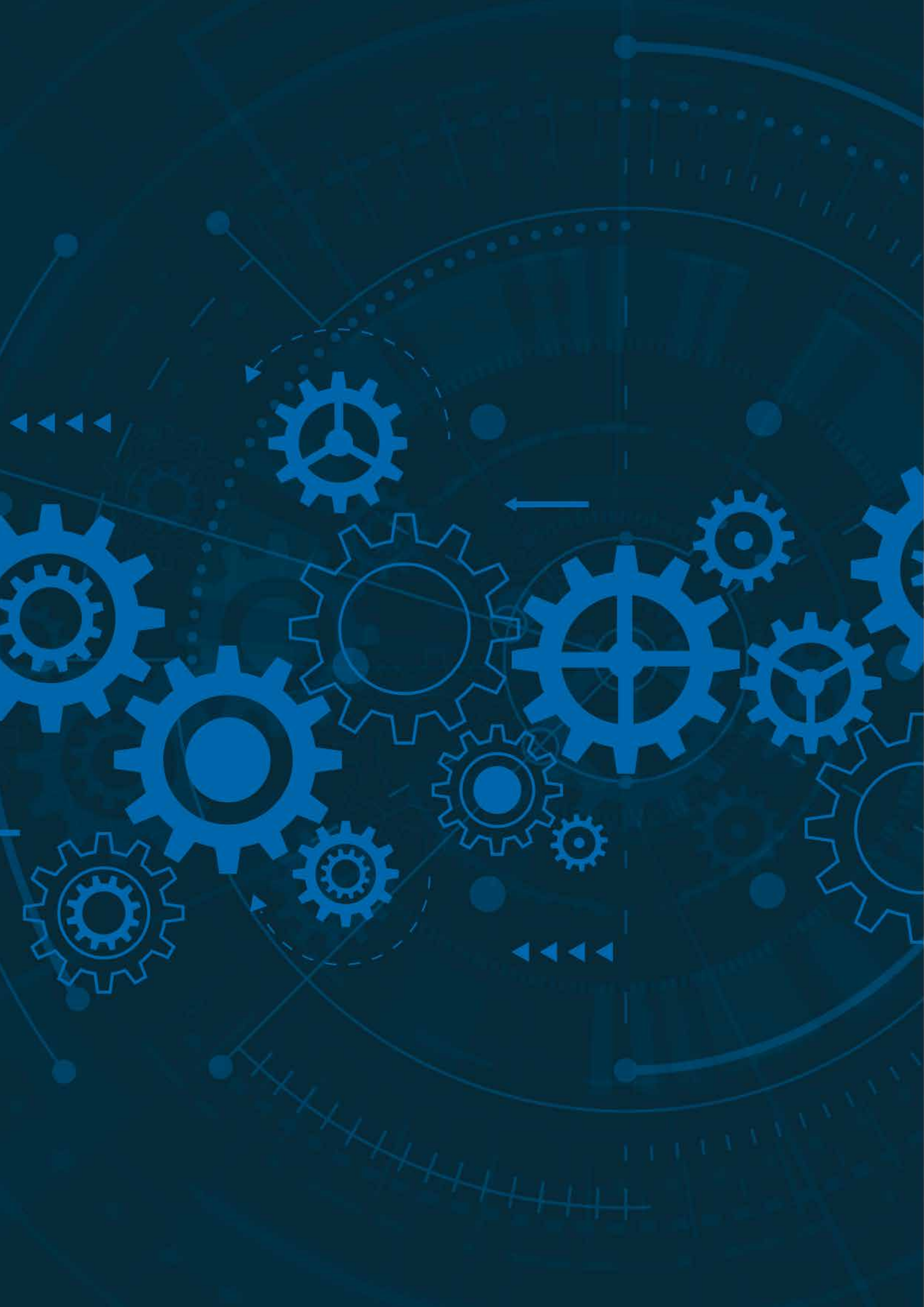
means striving to do good for people and the planet while minimizing harms arising from evaluation as an intervention.

- Explicit and ongoing consideration of risks and benefits from evaluation processes, products and longer-term consequences.
- Maximizing benefits at systemic (including environmental), organizational and programmatic levels.
- Doing no harm and not proceeding with an evaluation when harms cannot be mitigated.
- Ensuring evaluation makes an overall positive contribution to human and natural systems and to the mission of the United Nations.

Source: United Nations Evaluation Group, 2020: Ethical Guidelines for Evaluation.

Chapter II

IFAD's evaluation and learning system



II. IFAD's evaluation and learning system

25. The Revised IFAD Evaluation Policy establishes the broad framework through which evaluative evidence is produced and used. The policy emphasizes the need for effective use of and learning from evaluation products. Similarly, the use of evidence as the basis for decisions on the design and implementation of projects, programmes and strategies is at the core of the IFAD Strategic Framework 2016-2025,[22] Development Effectiveness Framework, IFAD11 and IFAD12. This chapter presents IFAD's evaluation architecture, its components, functions and the types of evaluations undertaken. It also identifies key processes for knowledge generation and evidence use.

A. IFAD's evaluation architecture

26. IFAD's evaluation architecture consists of the independent and self-evaluation systems, which provide important tools for accountability, learning and knowledge management, with useful practical applications at both the strategic and operational levels.

27. Figure 4 shows the IFAD evaluation architecture, which combines independent and self-evaluation, as well as linkages to development partners and IFAD's ultimate clients – small-scale rural producers and their communities.

28. Independent evaluations are conducted by IOE, which is structurally, functionally and behaviourally independent of Management. From a governance standpoint, the IFAD Revised Evaluation Policy states that IOE reports directly to IFAD's Executive Board and that the Evaluation Committee supports the Executive Board in evaluation matters. IOE ensures that the entire evaluation function at IFAD follows internationally recognized good standards and practices. Independent evaluations help to reveal what has been achieved, what does or does not work and why, and to guide the development of successful policies and strategies to support rural transformation. The target audiences of independent evaluations are IFAD's Management and governing bodies, Member States and the international development community at large.

29. The self-evaluation system is the responsibility of Management and is overseen by the Operational Policy and Results Division (OPR) and the Research and Impact Assessment Division (RIA).[23]

30. Self-evaluation serves three primary purposes: (i) to obtain real-time feedback on performance and inform decision-making; (ii) to learn from experience and improve the development effectiveness of operations; and (iii) to report to IFAD's governing bodies on aggregate results against targets agreed upon with Members, as well as the results and impact attributable to its operations. Activities related to the first objective are carried out by project teams in regional divisions of the Programme Management Department, while OPR provides guidance and support. Activities related to the second objective are led by OPR and the Strategy and Knowledge Department (SKD) and involve project teams. The third objective is the responsibility of OPR (results) and RIA (impact). Beyond providing robust results reporting for accountability, self-evaluation products are one of the sources of information for independent evaluations.

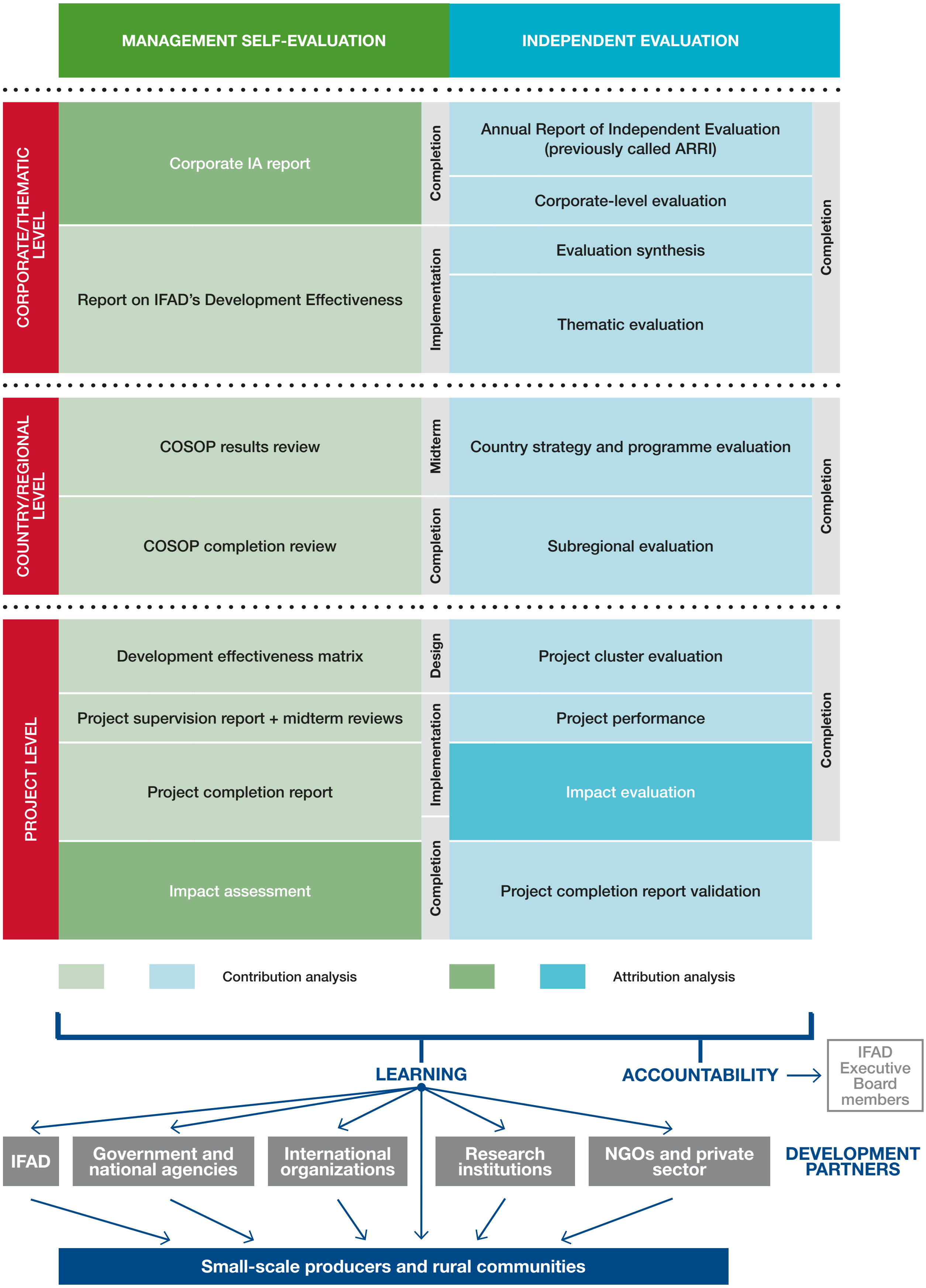
31. Strong monitoring and evaluation (M&E) systems, which promote adaptive management and learning, or monitoring, evaluation, adaptation and learning (MEAL), are of paramount importance. In addition, the impact assessments conducted by RIA that report on attributable impacts of IFAD's investments on its goal and strategic objectives make IFAD the only IFI with this type of systematic and rigorous approach to corporate reporting. High-quality self-evaluation products are critical for effective IFAD-wide evaluations, as well as evidence-based programming.

22 IFAD Strategic Framework 2016-2025.

23 Within IFAD's structure, OPR is part of the Programme Management Department, while RIA is part of SKD.

FIGURE 4

A simplified scheme of IFAD's evaluation architecture (up to IFAD11)



32. Both independent and self-evaluation functions work at three different levels: (i) project; (ii) country or regional; and (iii) corporate or thematic. The details of each evaluation product are provided in part 2 of this manual. The first level of project evaluation provides the basis for the analysis of the other levels. Self-evaluation is conducted at the design, implementation and completion stages of the project. Furthermore, the majority of evaluations are based on contribution analysis, which aims to measure interventions' contribution to the overall change. To complement these analyses, impact assessments or evaluations, conducted by RIA and IOE, respectively, and corporate impact assessment reporting, conducted by RIA, are based on attribution analysis. They evaluate the impacts attributable to the interventions isolated from all factors that might have affected the outcome at both the project and corporate levels. Chapter III presents a thorough discussion on contribution and attribution analyses.

33. The entire evaluation architecture provides evidence for accountability, learning and knowledge internally at IFAD for strategic and operational purposes, as well as externally for development partners and, eventually, end clients. Development partners are governments and national agencies, international organizations (e.g. other United Nations organizations), research institutions, non-governmental organizations and the private sector. All evaluation products are publicly available to contribute to evidence-based programming by all partners. IFAD's end clients are small-scale producers and rural communities, which directly benefit from evidence generated by IFAD evaluation functions through improved project design, and indirectly, through improved evidence available to all development partners. Annex I, in part 1 of this manual, contains a full description of the evaluation system, including its components, and the linkages and learning loops between the independent and self-evaluation systems.

B. Enhancing learning and evidence use

34. The ultimate purpose of evaluation is to inform decision-making and contribute to a broader knowledge base within IFAD and among IFAD's external stakeholders, such as national and international development partners. A thorough and rigorous evaluation and the production of a good report are not enough for an evaluation to be useful.

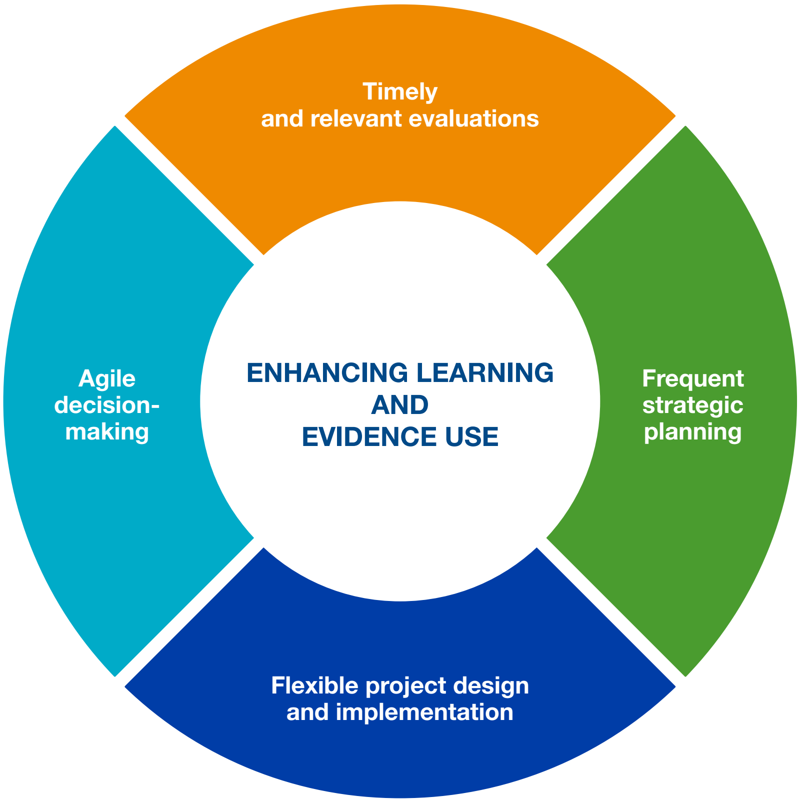
35. Consistent with international practice,[24] IFAD aims to maximize the use of evidence and evaluations throughout the evaluation process. Adopting an adaptive, learning-centred approach requires IFAD to learn and respond quickly and effectively to evidence and lessons. As illustrated in figure 5, the generation and use of evidence and learning must be a continuous process to ensure that IFAD becomes more agile, responsive, innovative and effective in the solutions it offers. IFAD needs to: (i) generate timely and relevant evidence; (ii) foster dialogue and strategic planning with development partners; and (iii) ensure flexible project design and implementation that constantly react to emerging results and learning, as well as external changes and events.

36. The use of evidence for quality decision-making is also at the core of the updated Development Effectiveness Framework (DEF 2.0).[25] In DEF 2.0, all objectives and activities proposed to enhance existing self-evaluation tools have a strong focus on learning (as a cross-cutting principle) and the necessary incentives, tools, mechanisms and approaches to ensure that learning is prioritized and prized by IFAD's staff, government partners and beneficiaries. Management is working to ensure a strengthened culture of learning in the coming years.

24 See, for example: World Bank (2019) World Bank Group Evaluation Principles.
25 https://webapps.ifad.org/members/ec/115/docs/EC-2021-115-W-P-6.pdf

FIGURE 5

Learning and evidence use – an iterative process



37. The organizational and functional independence of IOE is essential for ensuring sound, credible and transparent evidence consistent with international norms and standards and the principles set forth in the Revised IFAD Evaluation Policy. At the same time, leveraging all components of the evaluation system can generate a virtuous learning circle for IFAD. Thus, collaboration between IOE and Management is key to ensuring the relevance, timeliness and utility of the evaluations and strengthening that virtuous learning loop. In this regard, the pathways and processes outlined below do not hamper IOE's independence and are designed to introduce innovative elements in the evaluation processes and learning loops to complement the existing ones.

38. Drawing from and contributing to IFAD's Knowledge Management Strategy (box 2), synergies and complementarities between independent and self-evaluations are maximized in two broad action pathways:

- Evidence and knowledge generation; and
- Knowledge use within and beyond IFAD.

39. Together, these pathways are intended to lead to the creation of an evidence base useful to both IFAD and its partners and systematically embedded in IFAD's operations.

BOX 2

IFAD Knowledge Management Strategy (2019)

IFAD's Knowledge Management Strategy for the period 2019-2025 is part of the Fund's approach to increasing its development effectiveness. The strategy and its accompanying three-year action plan will help IFAD leverage the best available and most relevant knowledge, based on both evidence and practice, from its own work, work with partners, and other external sources.

Activities will be implemented in three broad areas: knowledge generation; knowledge use; and building the enabling institutional environment for evidence-based learning and knowledge sharing.

Source: https://www.ifad.org/en/-/document/knowledge-management-strategy

40. Different evaluations require different levels of engagement at different points in time. In addition, implementation of specific processes and activities may vary with the type of evaluation and the needs and contexts of specific institutional operating and client environments, as discussed in part 2 of this manual. Yet, there are overarching approaches that lay the foundations for knowledge generation and use, as described in pathways 1 and 2.

Pathway 1: Evidence and knowledge generation – IOE-Management collaboration

41. To maximize the synergies between independent and self-evaluations, collaboration between IOE and Management is pursued throughout the evaluation process, consistent with the independence of IOE. This includes the processes for selecting, planning and designing evaluation products, conducting evaluations and ensuring the identification and sharing of lessons learned and recommendations (box 3).

BOX 3

Examples of cooperation between IOE and Management throughout an evaluation process

Evaluation selection, planning and design

- IOE and Management cooperate to identify evaluation needs and demands. Independent evaluations and thematic-oriented self-evaluations will be planned, based on the strategic and operational needs of IFAD and external stakeholders. For example, IOE and Management collaborate to identify evaluation topics.
- IOE and Management collaborate to prepare a multi-year strategy and annual evaluation plans (self and independent) for Executive Board approval, and to heighten the relevance and improve the timing of different evaluation products, for example, between country strategic opportunities programmes (COSOPs), COSOP completion reviews and country strategy and programme evaluations.

Conducting evaluations

- Synergies and collaboration are pursued throughout the conduct of evaluations. This includes exchanging information during midterm reviews, preparing other self-evaluation products and reducing overlaps that are not beneficial.

Evaluation recommendations and completion

- Joint technical reviews and learning events are undertaken before recommendations are finalized. The purpose of these meetings is to draw out and discuss lessons and recommendations to ensure full internalization and learning.
- IOE and Management jointly organize and hold round-table workshops and/or learning events to discuss the results and lessons emerging from the relevant evaluation with multiple partners. Similarly, learning workshop are organized for corporate-level evaluation at an appropriate stage in the process.

Pathway 2: Evidence and knowledge use within and across IFAD

42. The evaluation process does not end with the evaluation report. The follow-up steps are critical for ensuring knowledge use within and beyond IFAD and for making certain that the findings and lessons from each evaluation are communicated, absorbed and applied across institutions and shared with development partners and end clients in rural areas. To this end, key activities include (box 4):

BOX 4

Examples of use of evaluation evidence and knowledge

- **Synthesis of overarching findings**: Independent and self-evaluations are used to prepare syntheses of evaluations (including policy briefs and infographics) to inform relevant corporate policies, strategies and operational processes in IFAD.
- **Learning events**: IOE and Management organize capitalization workshops, both internal stock-taking events and events in collaboration with key development partners, as required, to discuss the findings of key evaluations to facilitate learning and the uptake of lessons.
- **Leveraging technology to learn from evaluation**: In order to enhance utilization, evidence from both self- and independent evaluations are easily accessible in appropriate formats for the target audience. Going forward, increasing the use of information and communications technology (ICT) can help systematize and extract lessons from all types of independent and self-evaluations to design more impactful development interventions.
- **Management action register**: A formal process in which IOE and Management discuss and identify action to be taken as a result of key recommendations from each major evaluation. Progress in this regard will be monitored by after-action reviews. Management plans to bring volume II of the President's Report on the Implementation Status of Evaluation Recommendations and Management Actions (PRISMA) online to create an online evaluation tracking system to record recommendations, identify management follow-up activities and report progress. The system will also allow data extraction for the quick generation of recommendation status reports. Similar electronic systems are in place in other multilateral organizations (e.g. United Nations Development Programme (UNDP), World Bank Group).
- **After-action reviews** of recommendations and their implementation for monitoring both performance and the results of the actions taken to address recommendations. After-action reviews may involve development partners as well.
- **Operational Strategy and Policy Guidance Committee – project and COSOP designs**: RIA's Impact Assessments create knowledge through attributable impact analyses that feed into project and COSOP designs to improve programming.

43. In conclusion, given the 2021 evaluation policy's emphasis on collaboration between IOE and IFAD Management, chapter II of this manual provides guidance on their interactions. Box 5 describes the key phases of an independent evaluation process and its interactions with IFAD Management. More details on the self- and independent evaluation steps and their interactions are presented in part 2 of the manual. It is important to underscore that other development partners and clients play a fundamental role in IFAD-funded operations and are key stakeholders in an evaluation process.

BOX 5

Overview of the phases of an independent evaluation process and its interactions with IFAD Management

Design

- **Draft approach paper**. IOE prepares an approach paper that outlines the objectives, scope, methodology and process of the evaluation. IOE interacts with the key partners, notably IFAD Management and the government (when applicable) and seeks their comments. For more complex evaluations, an inception workshop may be held and a preparatory mission conducted.
- **Final approach paper**. IOE finalizes the approach paper by addressing the comments of stakeholders and shares the paper with them.

Implementation

- **Field missions for data collection**. IOE plans the evaluation field missions in conjunction with IFAD Management, the government and other stakeholders, as required, and communicates with them in advance.
- **Wrap-up meetings**. When country visits are conducted, at the conclusion of the visits IOE organizes a debriefing with the government, IFAD operational staff and other stakeholders, as necessary.

Reporting

- **Draft report**. IOE prepares the draft evaluation report and shares it with IFAD, the government and other stakeholders, as necessary, for comments.
- **Findings and recommendations**. Insofar as possible, evaluation findings and recommendations are discussed internally and with stakeholders to enhance ownership and use. The purpose of these meetings is to elicit and discuss findings and recommendations to ensure full internalization and learning.
- **Management Response**. IFAD Management prepares a response to the evaluation, which is included in the final report and discussed at relevant sessions of the governing bodies along with IOE's comments.
- **Final report**. IOE finalizes the evaluation report by independently addressing the comments of IFAD and the government (and other stakeholders, as appropriate). The final report is shared with stakeholders together with an audit trail summarizing how the comments were addressed.

Completion and dissemination

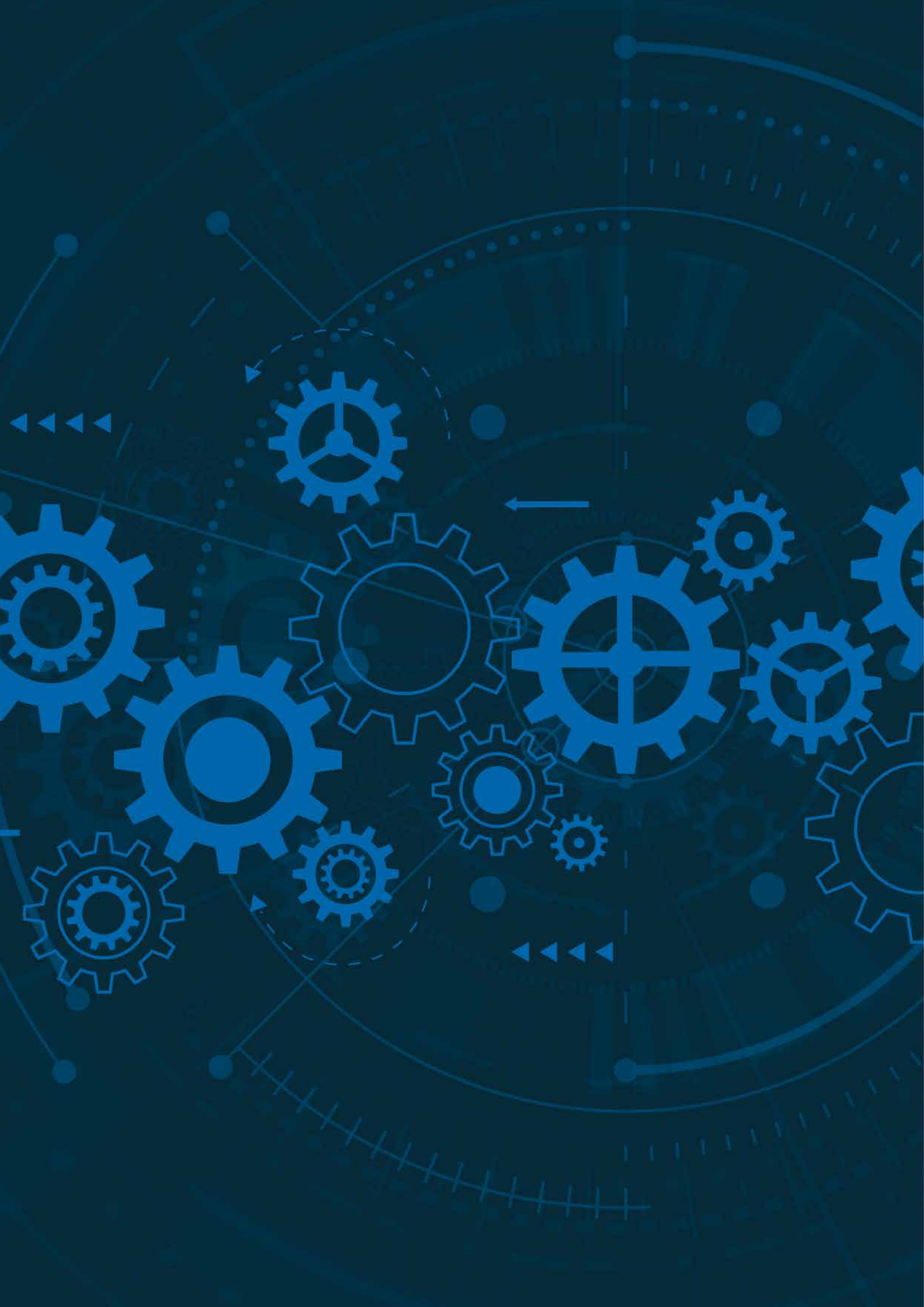
- **Final workshops**. IOE organizes final workshops and learning events in collaboration with IFAD Management, the government and other stakeholders (when applicable) to share and discuss the findings, lessons and recommendations.
- **Agreement at completion point**. For country-level evaluations, the agreement at completion point contains a summary of the evaluation findings and recommendations that IFAD Management and the government agree to adopt and implement within specific timeframes. IFAD Programme Management Department divisions are responsible for implementing the process. IOE takes note of the progress and final outcome.
- **Presentation to IFAD governing bodies**. Selected evaluations are presented to relevant sessions of the Evaluation Committee and Executive Board.
- **Dissemination**. IOE ensures appropriate dissemination of evaluations through the IOE website and newsletter. When required, IOE also collaborates with the core partners to disseminate the messages and evaluative innovations and methods.

44. IFAD evaluations play an important role in IFAD's knowledge management systems, as they generate and globally share knowledge on investing in sustainable and inclusive rural transformation. This, in turn, enables IFAD to play a greater advocacy role in supporting global efforts to achieve the SDGs.

45. Communication and dissemination of evidence and lessons beyond IFAD are therefore an integral part of the evaluation process. IFAD's evaluation policy clearly states that all independent and self-evaluation products shall be disclosed to the public and widely disseminated. At the design stage, all evaluation products include a communication and dissemination plan to ensure that evaluations are shared effectively. Specific dissemination approaches for each type of self- and independent evaluation product are described in part 2 of this manual.

Chapter III

Methodological fundamentals



III. Methodological fundamentals

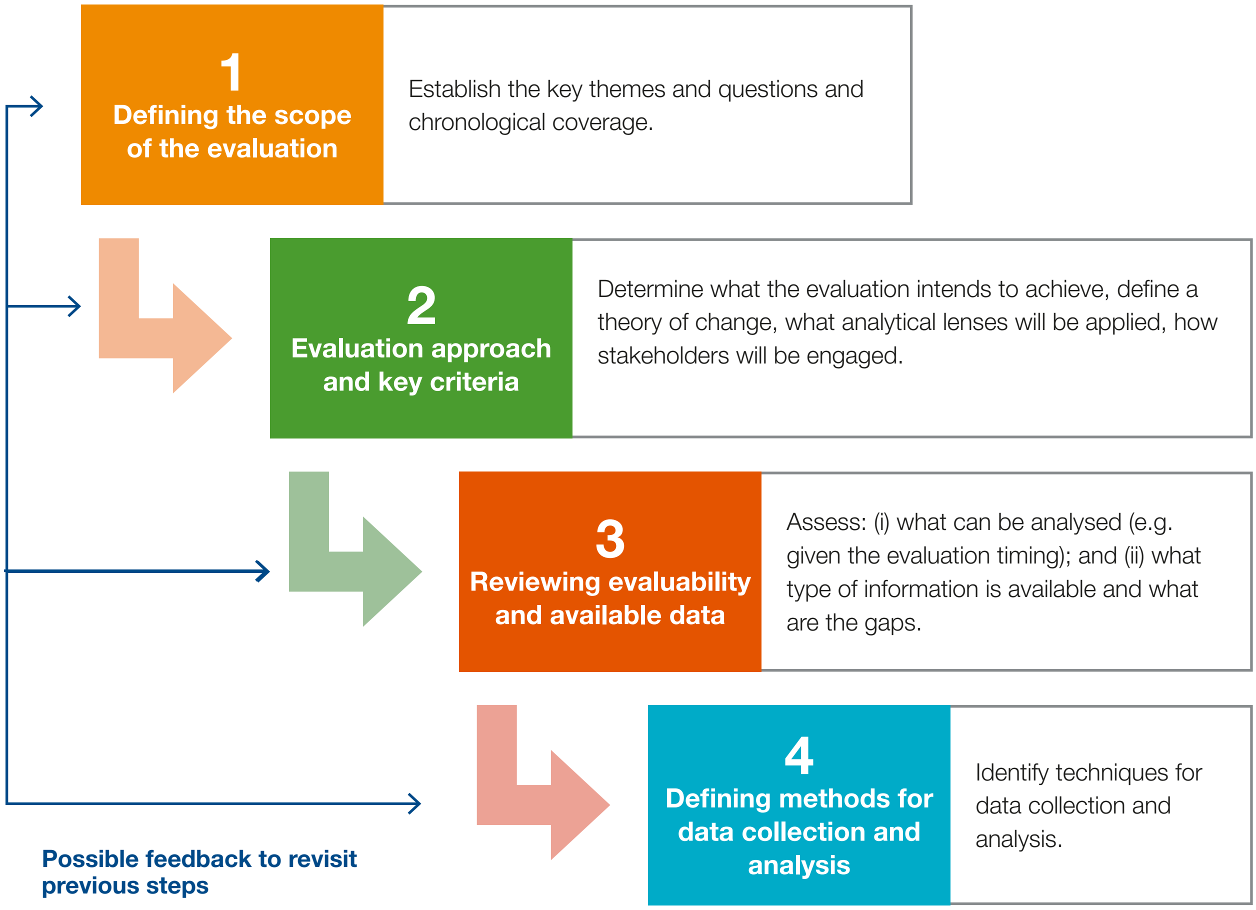
46. Chapter III is devoted to the methodological fundamentals for conducting evaluations at IFAD, ranging from scoping, design approaches, evaluation criteria and the ratings used to assess different types of performance to methods for data collection and analysis. It draws from contemporary literature on evaluation and the experience of IOE and RIA. Impact assessments conducted by RIA, which complement other evaluations with a unique approach, are based on theoretical and applied economics literature and are briefly introduced in a separate subsection. The chapter concludes with a section on tips for preparing conclusions and recommendations for evaluation reports.

47. While some sections (e.g. scoping and design approaches) apply primarily to independent evaluations, they also serve as a reference for other IFAD divisions to be used in line with the type and scope of evaluation conducted. The sections on evaluation criteria and ratings apply to the entire evaluation system. Evaluators' awareness of the methodological fundamentals helps to set standards and ensure consistency in methodology and reporting formats across evaluators and evaluations, enhancing the robustness and rigour of IFAD evaluation products. This chapter is not intended as a comprehensive guide to evaluation methodology but provides references to existing methodological guidance and studies. Specific issues and considerations for different types of evaluation products are covered in part 2 of this manual.

48. **Key steps in evaluation design**. The key steps presented in figure 6 are applicable to most evaluation exercises. They include: (i) definition of the scope (coverage) of the evaluation (topics, timeframe); (ii) adoption of an evaluation approach, elaboration of a ToC and evaluation criteria; (iii) the evaluability assessment (defining what can be evaluated and what data and information are already available); and (iv) the determination of specific methods for data collection and analysis. Figure 6 also shows possible feedback loops in the design steps – the evaluation steps can be considered iterative rather than strictly sequential. This chapter further elaborates on these steps.

FIGURE 6

Key steps in evaluation design



A. Defining the evaluand and the evaluation scope

49. While some evaluations (e.g. project- or country-level) can be standardized more easily than others, all evaluations should start by defining the evaluand (that is, what is being evaluated: for a thematic or strategic evaluation, this may require an operational definition) and the scope of the evaluation (e.g. the specific topics to be analysed, the time interval to be considered). This is particularly important for strategic evaluations and country-level evaluations, where, for example, the evaluators will need to determine the number of projects to be reviewed and the timeframe. Defining the scope and coverage may also be useful for project-level evaluations, where the analysis may need to focus on certain project components or specific topics.

50. To illustrate, the 2020 corporate-level evaluation of IFAD's support to innovations for inclusive and sustainable smallholder agriculture elaborated an operational definition of innovation: A new way of acting – practice, approach/method, process, product, or rule – brought in or implemented for the first time, considering the context, timeframe and stakeholders, with the purpose of improving performance and/or addressing challenge(s).[26] It also set overarching questions such as:

- To what extent (how and why) have corporate instruments, tools and approaches been successful in promoting agricultural innovations within country programmes?

26 https://www.ifad.org/en/web/ioe/-/corporate-level-evaluation-on-ifad-s-support-to-innovations-for-inclusive-and-sustainable-smallholder-agricultu-1

- To what extent (how and why) have IFAD's operations promoted agricultural innovations that: (i) have responded to smallholder farmers' needs/demand; and/or (ii) were targeted and inclusive?

- How have those innovations led to positive outcomes, and how have they been scaled up for sustainable and resilient development of smallholder agriculture?

51. Other examples of operational definitions that helped define the scope of strategic evaluations are presented in box 6.

BOX 6

Operational definitions in corporate evaluations and evaluation syntheses

Corporate-level Evaluation of IFAD's Financial Architecture – 2018

"Financial architecture" is defined as the policies and systems adopted to mobilize, manage, allocate and disburse financial resources to fulfil IFAD's mandate of helping to reduce rural poverty. The key elements of financial architecture can be summarized under four broad headings: (i) sources of funds; (ii) financial support instruments; (iii) allocation system; and (iv) management, oversight and governance. The financial architecture of IFAD is not a discrete "programme" or a "policy" underpinned by a dedicated logical framework. It is, rather, the result of the stratification of a number of policies and decisions made by governing bodies and Management over forty years.

Evaluation synthesis on community-driven development in IFAD-supported projects – 2019

Community-driven development is a way to design and implement development policy and projects that facilitates access to social, human and physical capital assets for poor rural people by creating the conditions for:

- Transforming rural development agents from top-down planners into client-oriented service providers;
- Empowering rural communities to take initiative for their own socio-economic development (i.e. building on community assets);
- Enabling community-level organizations – especially those of poor rural people – to play a role in designing and implementing policies and programmes that affect their livelihoods; and
- Enhancing the impact of public expenditure on the local economy at the community level.

Source: Excerpted from: https://www.ifad.org/en/web/ioe/-/ifad-s-financial-architecture; and https://www.ifad.org/en/web/ioe/-/community-driven-development-in-ifad-supported-projects

B. Defining the evaluation design approach

52. An evaluation approach comprises the intended objectives of an evaluation and the analytical and axiological (i.e. the system of values) lenses to be applied. Different evaluation approaches have comparative advantages in addressing specific concerns and needs. Selecting the most appropriate evaluation approach is a vital stage in the overall evaluation process.

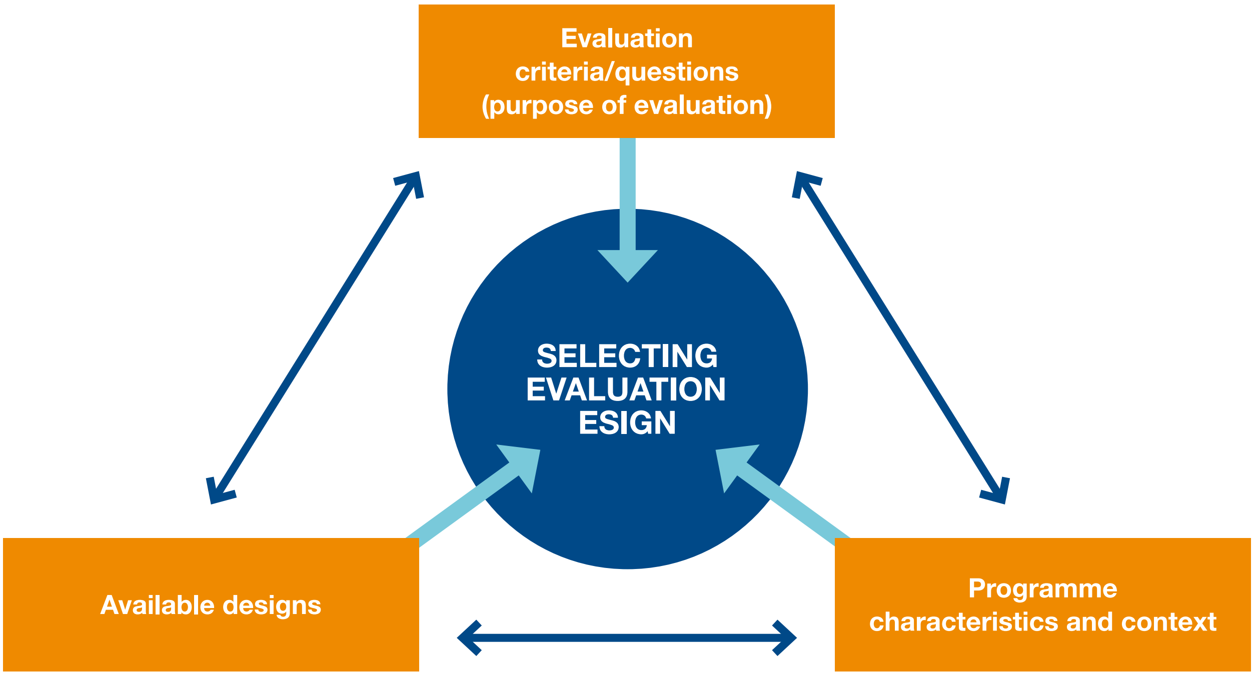
53. The evaluation design should match the purpose of the evaluation, the evaluation questions and the nature of the intervention. Figure 7 presents the Stern et al. (2012) Design Triangle[27] pinpointing the three factors that need to be taken into account when deciding on suitable evaluation designs: the evaluation questions that need to be answered; the "characteristics" of the intervention to be evaluated; and the range of available designs. For example, is the programme implemented in different settings, at different levels? A number of key considerations for evaluation design are specific to multilevel and multisite evaluations – e.g. country, regional, cluster evaluations – as this may involve analysing data at multiple levels of decision-making, sectors and locations. The methodology selected will enable the evaluation questions to be answered using credible evidence.[28] Throughout the following subsections, different methods options are presented to address different evaluation questions.

27 Stern, E., et al. *Broadening the range of designs and methods for impact evaluations. Report of a study commissioned by the Department for International Development*. Working Paper 38. (London: Department for International Development, 2021).

28 The Centre for Evaluating Complexity across the Nexus (CECAN) has recently launched a tool to support the process of methodological selection. See Befani, B (2020) Choosing appropriate evaluation methods. A tool for assessment and selection. CECAN. It also provides an overview of the potential and weaknesses of 15 methods.

FIGURE 7

The Design Triangle



Source: Stern et al. (2012), op.cit.

54. The following subsections present some recent approaches and practices that have emerged that evaluators may need to consider when defining the overarching evaluation approach, as required by the topic and context.

Emerging evaluation practices

55. Since the launch of the 2030 Agenda, there have been advancements in evaluation practices that are relevant to IFAD. Box 7 provides a brief summary of four practices stemming from the 2030 Agenda and the Sustainable Development Goals. The first two advancements, "transformative change," and "complexity and systems thinking," are of particular importance when evaluating large-scale or global phenomena, processes and systems.

56. The third advancement relates to how evaluations address the sustainability of interventions in the context of climate change, as this has far-reaching consequences for the environment and society.

57. The fourth advancement builds on the "no one left behind" principle of the 2030 Agenda. This requires evaluation efforts to integrate social justice, gender intersectionality and culturally responsive principles, which aligns well with IFAD's value of equity and focus on the poorest people and communities. Annex II provides additional details on social justice and gender-responsive evaluations at IFAD.

BOX 7

SDGs and their implications for evaluation practice[29]

Transformative change: The 2030 Agenda recognizes the need to transform societies through sustainable, resilient and inclusive paths. There is general agreement that achieving the SDGs requires transformational changes at scale that address root causes and systemic drivers of poverty, exclusion and pollution. Evaluators are increasingly called upon to answer questions about effects on norms and systems. In order to understand how interventions, programmes and policies support lasting system change, evaluators need to engage with strategic and aggregate-level evaluations and understand how systems-level change can be evaluated.

Complexity and systems thinking: The solutions to poverty, inequality, climate change and other global challenges are deeply interrelated in complex ways. For example, IFAD's ambition to contribute to SDG 2 couples natural processes with social and economic processes. Understanding these interactions requires the use of more sophisticated evaluation methodologies that include complexity science and systems analysis to assess the interconnectedness and trade-offs. The focus on complexity also stresses the need to focus on the context and analyse how programme outcomes are influenced by economic, political, sociocultural, ecological and other factors in the local, national and international context. A paradigm shift is therefore emerging, moving from the linear, hierarchical and static logframe to a more complex and dynamic approach to examine whether and how outcomes and impact are achieved. The application of theory-based approaches to evaluation offers valuable methods for design, data collection and analysis of findings.

Sustainability and climate resilience: The SDGs identify climate change not only as one of its specific objectives but as a threat multiplier with the potential to undermine progress on poverty, hunger, equality and health. Evaluating sustainability and resilience requires a different methodological approach, not the conventional evaluations of programme outputs and outcomes. Both sustainability and resilience involve assessing the ability of communities or other entities to respond to, cope with and adapt to shocks and stresses, which may occur over extended periods and are usually unpredictable. Evaluation efforts are seen as essential for providing evidence on whether actions to address the complex challenges associated with climate change are on track for achievement of the SDGs. To this end, evaluators must engage with climate data and analysis in the evaluation of interventions and policies.

Social justice and gender intersectionality – No one left behind: The commitment to leave no one behind is at the heart of the 2030 Agenda. Gender equality, reducing inequalities and ensuring that "no one is left behind" are considered distinct but linked core principles and goals of the SDGs. In line with the UNEG guidance on integrating human rights and gender equality in evaluation and practice,[30] evaluations must consider all these dimensions, as well as their intersections (for example, intersection of gender and other socio-economic characteristics, such as ethnicity, age, caste and income level).

UN Women defines intersectionality as: overlapping, concurrent forms of oppression which point to the depths of inequalities and the relationships among them in any given context. Using an intersectional lens also means recognizing the historical contexts surrounding an issue.[31] Intersectionality is about:

- Fighting discrimination, **within** discrimination;
- Tackling inequalities **within** inequalities;
- Protecting minorities **within** minorities.

Evaluations must also incorporate these principles throughout the process in order to address multiple causes of discrimination and exclusion (e.g. age, race and ethnicity, social status, disability) and the way they interact in a specific context. Participatory methodologies, unpacking the assumptions of the power relations in evaluations, offer approaches to conduct more inclusive evaluations.[32,33] By using participatory methodologies, evaluators will work in ways that fully consider differential impacts by gender and the way they interact with other forms of discrimination.[34,35]

29 The information in the box has been obtained from several sources: IIED, *Five considerations for national evaluation agendas informed by the SDGs* (London: IIED, 2016).
Rob D. van den Berg, Indran Naidoo, and Susan D. Tamondong, eds., *Evaluation for Agenda 2030: Providing Evidence on Progress and Sustainability* (Exeter, UK: IDEAS, 2017).
J. I. Uitto, Geeta, B. (eds.). Transformational Change for People and the Planet. *Evaluating Environment and Development. Sustainable Development Goals Series*. Springer (2022).
J.I. Uitto, J. Puri, and R.D. van den Berg, *Evaluating Climate Change Action for Sustainable Development: Introduction* (2017). In: Uitto J., Puri J., van den Berg R., eds. (2017) Evaluating Climate Change Action for Sustainable Development. Springer, Cham.
Michael Bamberger, Marco Segone and Florencia Tateossian *Evaluating the Sustainable Development Goals With a "No one left behind" lens through equity-focused and gender-responsive evaluations*. (New York: UN Women, 2016).
Rob D. van den Berg, Cristina Magro, and Silvia Salinas Mulder (eds), *Evaluation for Transformational Change: Opportunities and Challenges for the Sustainable Development Goals*. (Exeter, UK: IDEAS, 2019). See also as a reference: Patton, M.Q., Blue Marble *Evaluation: Premises and Principles*, (New York: Guilford Publishing, 2019).

30 The UNEG Handbook provides a tool for integrating human rights and gender equality into evaluation practice. United Nations Evaluation Group, *Integrating Human Rights and Gender Equality in Evaluations* (2014), **http://www.unevaluation.org/document/detail/1616**

31 UNEG, *Integrating Human Rights and Gender Equality in Evaluation and Practice: Towards UNEG Guidance* (New York: UNEG, 2011).

32 A. Stephens, E.D. Lewis and S.M. Reddy, *Inclusive Systemic Evaluation (ISE4GEMs): A New Approach for the SDG Era (New York: UN Women*, 2018).

33 Evaluation Cooperation Group, "Gender. Main messages and findings from the ECG Gender practitioner's workshops" (Washington, D.C., 2017).

34 A. Stephens, E.D. Lewis and S.M. Reddy, *Inclusive Systemic Evaluation (ISE4GEMs): A New Approach for the SDG Era* (New York: UN Women, 2018) **https://www.unwomen.org/en/news/events/2018/10/ise4gems-launch-event**

35 Evaluation Cooperation Group, "Gender. Main messages and findings from the ECG Gender practitioners' workshops" (Washington, D.C., 2017), **https://www.ecgnet.org/document/main-messages-and-findings-ieg-gender-practitioners-workshop**

Applying theory-based approaches

58. In line with current international practice, IFAD evaluations follow a theory-based approach.[36] While different definitions of theory-based evaluations exist, they are based on an explicit ToC that explains the theory of a development intervention. The evaluation is then designed to test the theory.[37] Theory-based approaches, therefore, contrast with evaluation approaches that look solely at outcomes. Furthermore, theory-based evaluation is part of an approach to evaluation and not a specific method or technique. It is a way of structuring and undertaking analysis in an evaluation, which helps to establish whether the linkages between interventions and intended impacts are plausible, account for other contributory factors and capture unintended effects.

59. Theory-based approaches seek to: (i) identify and explain the influence of context on programme results; (ii) understand the underlying operating mechanisms that generate the observed effects and how these effects vary from context to context and population to population. Thus, theory-based approaches go beyond assessing "what has changed" to answer the more difficult questions of how, why, where and for whom as well.

60. Importantly, theory-based approaches highlight the assumptions, conditions and risks that sometimes are left as implicit in the design of a project, strategy and policy. In doing so, they help identify possible gaps in the logical chain.

61. In some cases, a project, country or corporate strategy and policy are elaborated using a ToC. When a ToC has not been explicitly developed, the evaluation teams may reconstruct, making it explicit that they will have to elaborate one. In such cases, it is important to seek feedback from the main stakeholders to ensure that the reconstruction is realistic and reflects stakeholders' understanding.

62. Two important practical tips for developing a ToC are: (i) it should be relatively simple to understand and represent graphically: if it cannot be explained, it can hardly be understood, and, therefore, will not be used; (ii) it is a living instrument that must be revisited during the evaluation process – benefiting from stakeholders' insights.

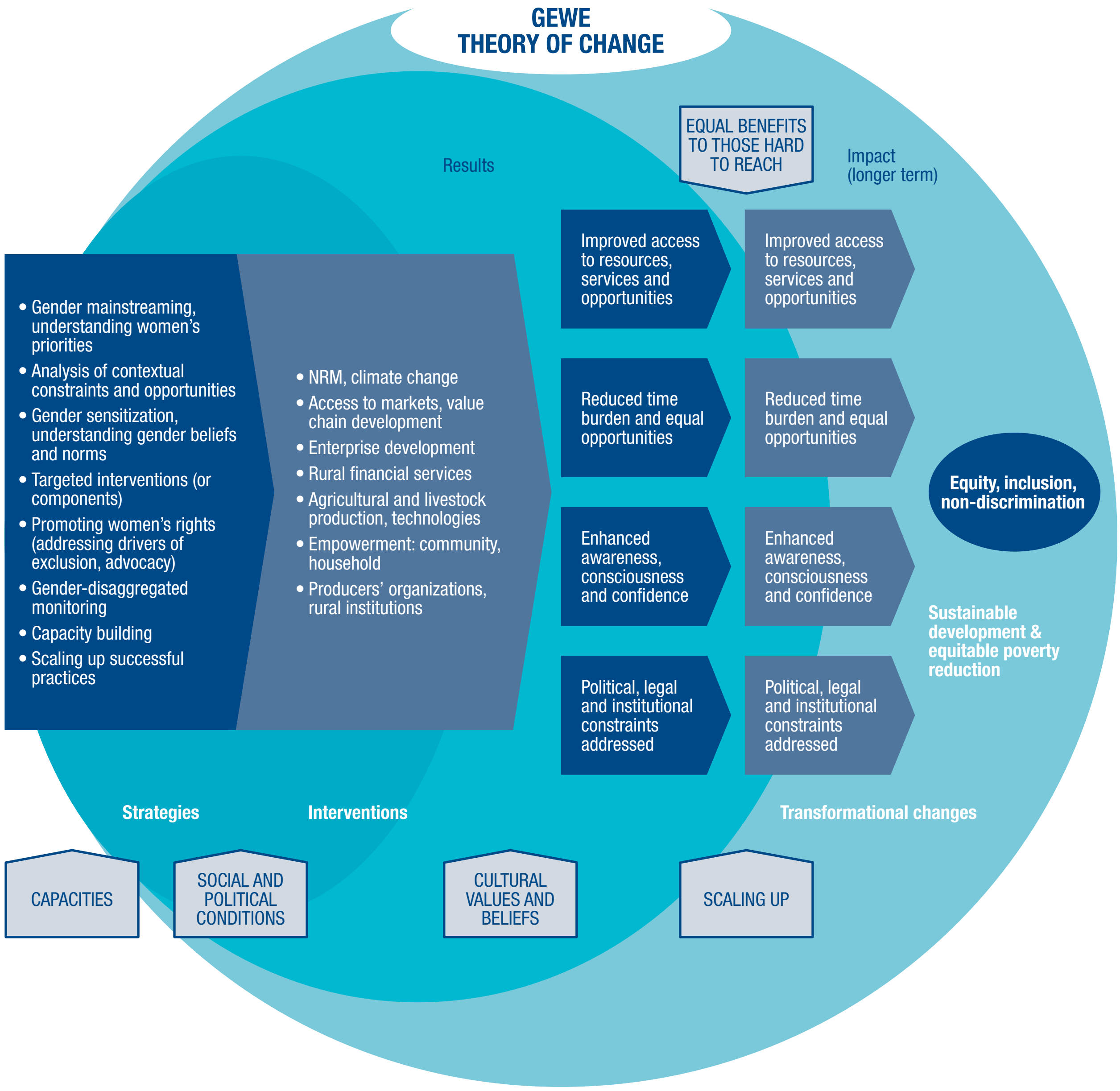
63. Examples of ToCs are presented in figure 8 (for an evaluation synthesis prepared by IOE) and figure 9 (for the impact assessment of a project by RIA).

36 Evaluation Cooperation Group, *ECG Big Book on Good Practice Standards.* (Washington D.C., 2012).

37 Policy interventions (programmes and projects) rely on underlying theories regarding how they are intended to work and contribute to processes of change. These theories, usually called theories of change, are often explicitly described in documents but sometimes exist only in the minds of stakeholders. Programme theories (whether explicit or implicit) guide the design and implementation of policy interventions and also constitute an important basis for evaluation. See for example: Jos Vaessen, Sebastian Lemire, and Barbara Befani. Evaluation of International Development Interventions: *An Overview of Approaches and Methods.* (Washington, D.C.: Independent Evaluation Group, World Bank, 2020).

FIGURE 8

Theory of change of the 2017 Evaluation Synthesis on Gender Equality and Women's Empowerment (GEWE)



Note: NRM = natural resource management.

FIGURE 9

Theory of change from the Impact Assessment of Peru: Strengthening Local Development in the Highland and High Rainforest Areas Project

INPUTS AND ACTIVITIES

Establishment of producer association

- Accompanied by elaboration of business plan
- Public competitions among PAs
- Formal registration of Pas
- Inclusion of women

Producer capacity building

- Training on production technologies
- Training on business management
- Training through classes, fellowships, local champions and constant accompaniment through local technical experts

Financial education

- Financial education workshops
- Access to financial instruments for women (savings accounts and life insurance)

OUTPUTS

- Producer associations are established and functioning with greater membership, including of women
- Producers are trained on production technologies and techniques
- Improved sheds/ ponds/etc. are built
- Producers are provided with information and skills on business management
- Producers and their household members receive financial education
- Women are provided access to savings and insurance instruments

OUTCOMES

Household level

- Adoption of new technologies
- Increased productivity
- Increased market participation
- Increased savings
- Improved business management capacities
- Increased group participation

Producer association level

- Adequate financial planning
- Improved PA technology and infrastructure
- Technical service provision capacity to members
- Increased involvement of women in PAs
- Increased cooperation among members

IMPACTS

Household level

- Increased income
- Increased food security/nutrition
- Increased resilience
- Empowerment of women

Producer association level

- Ability to maintain and expand activities
- Ability to access external finance
- Ability to grow and become a cooperative

ASSUMPTIONS

Inputs and activities:

- There is sufficient demand to form associations
- There is sufficient demand for training
- Cultural norms allow women to participate

Outputs:

- Training is appropriate and will lead to adoption
- Inputs are available to take advantage of new skills
- Technologies are adequate for local circumstances
- Members are willing to participate and contribute to association

Household level

- Markets for inputs, credit, output, etc. exist and function well
- Producers face no other barriers to improving productivity such as soil quality, capital, weather conditions, etc.
- Change in mentality towards women’s role in society

Producer association level

- Sufficient ability to access markets
- Reasonable support to PAs by local government
- Women can access leadership roles

64. Box 8 refers to two theory-based paradigms that are often used or referred to in the literature and help to operationalize theories of change. Box 8 captures two prominent categories: realist evaluation and contribution analysis.[38] Both perspectives were recently featured in guidelines by the ECG for gender-responsive evaluations.[39,40] These perspectives can be used in combination with several data collection methods [41] and techniques (chapter III.E). For example, a case-based method can be incorporated with a theory-based design, assessing a number of different interventions, each as a separate case, and using a method such as contribution analysis (see box 8) to assess causality for each one.[42]

38 For a detailed practical example of how to implement and maximize the value of contribution analysis see, for example: CDI *Contribution Analysis and Estimating the Size of Effects: Can We Reconcile the Possible with the Impossible?* Practice Paper (East Sussex: Centre for Development Impact, 2019).

39 Evaluation Cooperation Group, *Integrating gender into project-level evaluation* (ECG reference document (Washington D.C., 2017).

40 Note that these are examples of theory-based applications. There are other theory-based evaluation approaches, for instance, in impact evaluations – for example, the 3ie portal **here**.

41 An illustrative practical example of how combined methods and designs can be found in IFAD's pilot Participatory Impact Assessment and Learning Approach evaluation (PIALA), which combined a theory-based, mixed-methods approach to evaluation that was essentially participatory.

42 For detailed step-by-step guidelines on theory-based evaluations, see, for example: *BEAM Exchange Guidelines for evaluating the impact of market systems* (the BEAM Exchange is a platform for knowledge exchange and learning about the role of markets in poverty reduction).

BOX 8

Theory-based evaluation paradigms: two examples[43]

Realist evaluation: Primarily designed to answer the questions of: what works, for whom, in what respects, to what extent, in what contexts and how? The basic message of realist evaluation is that evaluation needs to focus on understanding what works better for whom and under what circumstances, and what aspect within a programme makes it work. Under a realist perspective, an evaluator would typically investigate the underlying mechanisms that generate the effects, how they interact with the context and main stakeholders, leading to differentiated results. They are most appropriate for evaluating new initiatives that seem to work but "where, how and for whom" is not yet understood; and/or for programmes that will be scaled up, to understand how to adapt the intervention to new contexts.

Examples:

- IFAD's FoodTrade East and Southern Africa Regional Programme Final Evaluation (2019). This evaluation combines realist approaches enquiry with contribution analysis, case studies and thematic synthesis.
- Investment Climate Fund (2020). Portfolio Evaluation. Support for Policy Change. This evaluation used a realist approach using qualitative data analysis software.

Contribution analysis (CA): Introduced by John Mayne (see Mayne 2012), this is primarily designed to answer the question: how much did the programme contribute to change? CA is a theory-based approach for analysing causality and is used alongside a ToC that explicitly describes how change is, or was, supposed to happen. It is essentially a narrative approach that can be supported by various types of evidence, where the evaluator formulates and then tests a contribution story that explains how the intervention has supposedly achieved (or is supposed to achieve) its impact. The contribution story is usually visualized as a causal chain of intermediate steps or outcomes, with assumptions and risks that make each step more or less likely to materialize. CA is particularly useful in fields such as research, policy influencing, markets, capacity building and mobilization, where there are often many different contributors to change. A distinguishing feature of CA is the emphasis on identifying plausible alternative explanations to observed results. These could include, for example, other related government programmes and economic or social trends.

Examples:

- Contribution analysis of a Bolivian innovation grant fund: mixing methods to verify relevance, efficiency and effectiveness (Giel Ton, 2017).

43 HM Treasury, *Magenta Book Analytical methods for use within in evaluation*. (United Kingdom, 2020)
J. Mayne, (2012) 'Contribution analysis: Coming of age?', *Evaluation*, 18(3), 270-280.
For a discussion of the practical implications of realist and contribution analysis approaches for large, complex, multi-country complex interventions, see: Centre for Development Impact (2020) Reality Bites: Making Realist Evaluation Useful in the Real World, CDI Practice Paper 22, and (2019) *Contribution Analysis and Estimating the Size of Effects: Can We Reconcile the Possible with the Impossible?*, CDI Practice Paper 20.

Systems mapping as a tool to support theory-based evaluations

65. The use of ToC is important in most evaluations, but a number of key considerations for evaluation design are specific to IFAD strategic and aggregate-level evaluations, such as country-level, thematic, project-cluster and corporate evaluations. These have the following characteristics: (i) they are multiproject evaluations that often cover multiple levels of interventions, multiple sites (communities, provinces, countries) and multiple stakeholder groups at different levels and sites; (ii) they contain both summative elements, as well as some formative focus and may contain important lessons for oversight bodies, Management, operations or other stakeholders.[44]

66. For this type of evaluation, the application of system mapping (see box 9) can be useful.[45] The visual aspect of system mapping puts complex concepts and relationships into simpler, pictorial representations.[46] An example of system mapping is presented in figure 10, showing a value chain as a system and presenting its subsystems and their boundaries (drawn from the 2019 Corporate level Evaluation on IFAD's Engagement in Pro-poor Value Chain Development). This mapping was instrumental to presenting the intricacy of working on value chain development, particularly when trying to generate equitable outcomes for smaller producers.

44 Hugh Waddington, Edoardo Masset, Emmanuel Jimenez, "What have we learned after ten years of systematic reviews in international development?" Journal of Development Effectiveness, (2018): 10:1, 1-16, DOI: 10.1080/19439342.2018.1441166.

45 Jos Vaessen, Sebastian Lemire, and Barbara Befani, *Evaluation of International Development Interventions: An Overview of Approaches and Methods*; Washington, D.C. Independent Evaluation Group World Bank, 2020.

46 For a practical example, see: Barbrook-Johnson, P. and Penn, A. *Participatory systems mapping for complex energy policy evaluation, Evaluation*, (2021): 27(1), 57-79.

BOX 9

Systems mapping

Systems mapping is particularly valuable for evaluating higher-level programmes characterized by significant causal complexity, providing a structured approach to identifying and presenting the systemic nature of programmes embedded in their contexts. The primary purpose of systems mapping is to describe the different components of a system and how they are connected. Mapping makes complex systems more comprehensible and therefore, more approachable. There are a number of ways to approach mapping the system to represent system elements and connections. They include: actor maps (to show which individuals and/or organizations are key players in the space and how they are connected); causal-loop diagrams (to clarify the positive and negative feedback loops that lead to system behaviour or functioning); issue maps (to lay out the political, social or economic issues affecting a given geography or constituency); mind maps (to highlight various trends in the external environment that influence the issue at hand); social network analysis (to understand the social structures and networks operating within the system); and many others.

In evaluation, system mapping is particularly relevant for understanding, for example, the institutional, social, economic and cultural aspects of the context in which a programme operates and how they influence how it works. This supports a better understanding of the nature and impact of complex programmes.

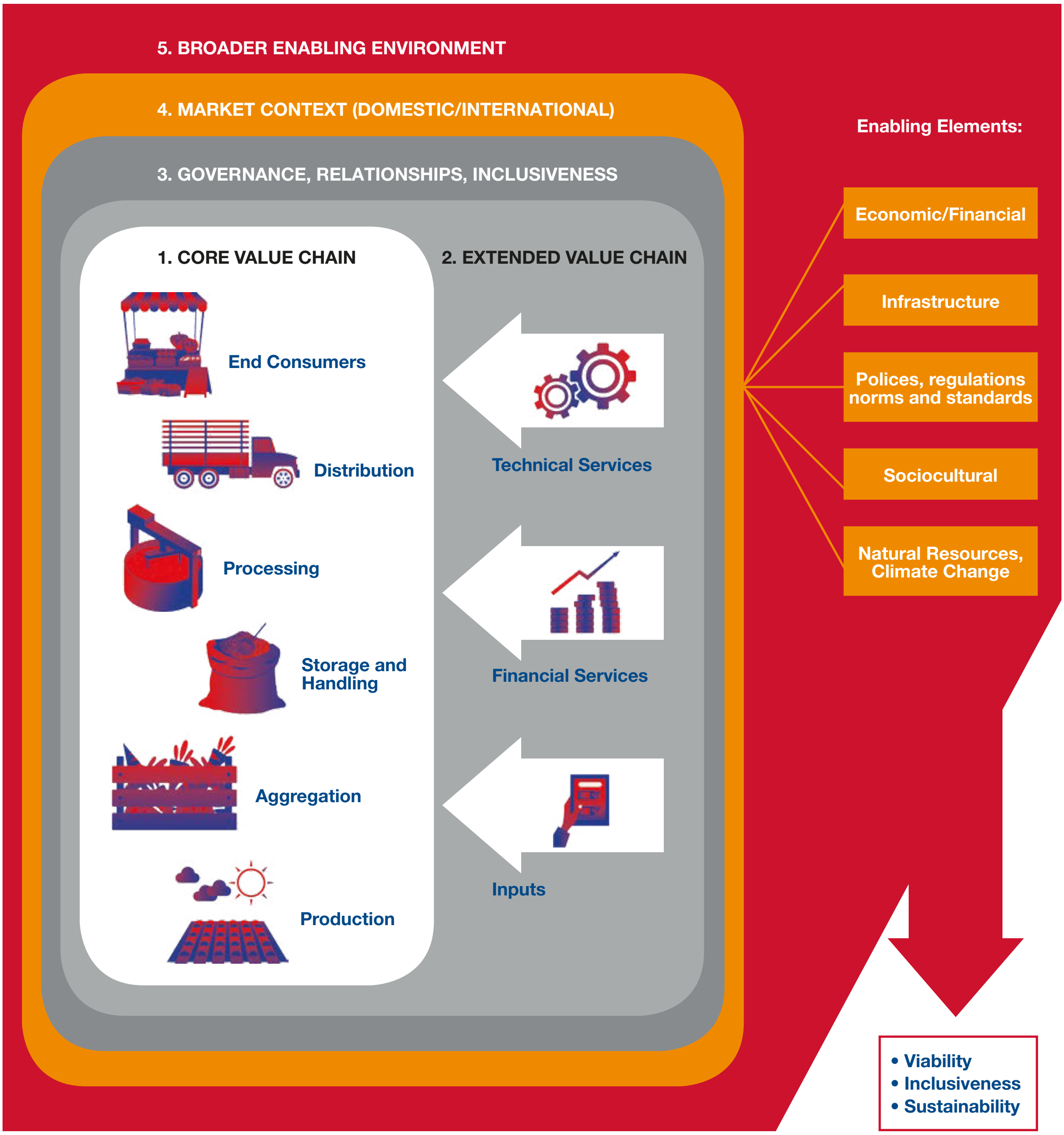
To ease the process, new online tools are also becoming available, several of which combine multiple mapping methods. Some of the more popular tools include: **Insight Maker** and **Kumu**.

Examples:

- Barbrook-Johnson, P. and Penn, A. (2021) Participatory systems mapping for complex energy policy evaluation, Evaluation, 27(1), 57-79.
- Participatory systems mapping: exploring and negotiating complexity in evaluation with BEIS and Defra. **https://www.cecan.ac.uk/wp-content/uploads/2020/09/1.4-Sys-Mapping-CECAN-Annual-Event-2018-AV2-P_1.pdf**

FIGURE 10

Mapping of a value chain system and its subsystems (Corporate-level Evaluation Value Chain, 2019)



Source: Corporate-level evaluation adaptation from the Food and Agriculture Organization of the United Nations (FAO) (2014), with inputs from GIZ (2018) and the United States Agency for International Development (USAID) (2014).

IFAD's attribution analysis approach

67. As highlighted in chapter I and chapter II, IFAD conducts evaluations assessing the contribution to overall change with the dual purpose of learning and accountability. Nonetheless, all evaluations are also aimed at detecting the extent to which a particular outcome or impact can be "attributed" to a given intervention. However, this is not a methodologically and operationally easy assessment. First, IFAD-supported activities involve many partners; therefore, disentangling the exact impact attributable to IFAD's intervention might not be straightforward. Second, beneficiaries are exposed to external factors that influence results, particularly donor countries' policies, beneficiary countries' domestic policies, other development programmes, socio-economic fluctuations, structural changes and climate phenomena. Third, baselines that could provide strong support for dealing with attribution issues are often unavailable or not of the required quality. Therefore, making a robust assessment attributing the results achieved on the ground to a particular intervention is challenging and expensive.

68. Given these challenges, theory-based design with a combination of methods for data collection, analysis and triangulation is the overarching approach to evaluation in IFAD, as discussed in greater detail in the next section. At the same time, IFAD has a unique approach to attribution analysis through both independent and self-evaluations.

69. IOE conducts impact assessments on selected completed operations, for the main purpose of: (i) establishing a more solid evidence base for future strategic evaluations; (ii) testing innovative methodologies for assessing the results of IFAD interventions more rigorously; and (iii) contributing to the ongoing internal and external debate on measurement of the impact of development interventions.

70. IFAD Management's systematic attribution analysis is conducted by RIA, which designs impact assessments specifically to detect the attributable impact of IFAD's interventions. This approach entails conducting impact assessments at the project level on a representative sample of 15 per cent of IFAD's operations portfolio. The data is then aggregated and projected to the entire portfolio at the corporate level, addressing the challenges to establishing attribution with a rigorous quantitative methodology. This process supplements the traditional independent and self-evaluation approaches with a systematic approach to attribution analysis. In addition, Management reports the results measured through the core outcome indicators (COI) in the project logframe,[47] which are also obtained through a rigorous survey methodology to establish attribution. This evaluation process makes IFAD unique among IFIs.

71. The quantitative methodology is based on economic theory and empirical literature on impact assessments using non-experimental ex-post methods.[48] It starts with the selection of a representative sample of projects closing in one IFAD replenishment period. Once the representativeness of the sample is confirmed, the methodology includes: (i) creating the projects' ToC; (ii) creating a robust sample frame that includes beneficiaries and a carefully selected control group (i.e. that represent the counterfactual); (iii) collecting quantitative and qualitative data from both groups (around 2,000 to 3,000 households); (iv) conducting analysis using non-experimental methodologies for each of the selected projects to estimate the attributable impact on a large set of development indicators (i.e. change in each indicator for beneficiaries compared to a control group); (v) conducting a meta-analysis to estimate overall impact; and (vi) conducting a projection analysis for the entire IFAD portfolio for corporate reporting and learning. In particular, the RIA Impact Assessments Report on the COI of IFAD's Strategic Framework 2016-2025 defined in the Results Management Framework of IFAD as Tier II development impact indicators. These are the economic goal of increasing incomes and the three strategic objectives of improving productive capacities and market access and strengthening environmental sustainability and climate resilience, as well as cross-cutting themes of food and nutrition security and women's empowerment.

47 The RIMS and associated logframe guidance was replaced in 2017 with CIs. In 2017, with the migration of all logframes from paper/PDF to ORMS, all RIMS indicators were converted to CIs in all logframes for ongoing projects. Core Outcome Indicators (COIs) are a subset of CIs and will become mandatory for all projects designed as of 2022. Results for COIs are captured through surveys conducted at three points over the course of project implementation: at project baseline, midterm and completion stages. A specific IFAD-tailored methodology has been developed for these surveys and is found in the core outcome indicators measurement guidelines.

48 For reference literature on the method used in this approach see: Rubin (1974), Rosenbaum and Rubin (1983), Heckman et al. (1997), Hahn (1998), Heckman et al. (1998), Dehejia and Wahaba (1999), Wooldridge (2007), Hirano et al. (2003), Caliendo and Kopeinig (2008), Imbens and Wooldridge (2009), Wooldridge (2010), Austin (2011).

C. Evaluation criteria, key questions and ratings

72. In line with good practice in international development evaluation, IFAD uses a set of evaluation criteria to assess the performance of policies, strategies, operations and business processes.[49] The use of evaluation criteria supports consistent, high-quality evaluation across IFAD and facilitates the aggregation of findings to conduct additional thematic analyses (e.g. across regions, topics, type of country), as well as analysis over time.

73. As shown in figure 11, in addition to the six internationally established criteria (OECD-DAC),[50] IFAD adopts further criteria that address its specific mandate.

74. This manual introduces four main changes with respect to previous editions. The purpose of these changes are: (i) to align with international standards; (ii) to avoid excessive fragmentation and repetition in the structure of reports, encouraging the preparation of more concise documents; and (iii) for Management to follow reporting practices agreed with Member States in the context of IFAD11 and IFAD12.

75. First, the manual introduces the "coherence" criterion (now part of the international reference criteria) to be used mainly for country and corporate/thematic evaluations.[51] "Non-lending activities" (knowledge management, partnership building and policy dialogue) are assessed as subdomains of coherence in evaluations at the country, corporate and thematic levels.

76. Second, the manual now merges the two criteria of "environment and natural resource management" and "climate change and adaptation" into a single criterion. This is in line with the Intergovernmental Panel on Climate Change (2018) definition of climate resilience: In human systems, the process of adjustment to actual or expected climate and its effects, in order to moderate harm or exploit beneficial opportunities. In natural systems, the process of adjustment to actual climate and its effects. In IFAD's criteria, adaptation is therefore not viewed in isolation. The key goals of adaptation strategies are to build the resilience of people and agricultural systems to climate change and to sustain and enhance the livelihoods of poor people. These strategies must consequently be rooted in an understanding of how the poor and vulnerable sustain their livelihoods, the role of natural resources in livelihood activities and the scope for adaptation action that reduces vulnerabilities and increases the resilience of poor people. For the purpose of self-evaluation, and only for projects closing until the IFAD12 period, the two dimensions of "environment and natural resource management" and "climate change and adaptation" will continue to receive separate ratings at the self-evaluation stage to allow reporting on RMF11 and RMF12 targets agreed with Member States. The two ratings will be aggregated for the purpose of IOE reporting.

77. Third, the manual presents a new arrangement of the criteria. "Innovation"[52] is assessed under effectiveness, whereas "scaling up" and "environment and natural resource management and climate change adaptation" are assessed under sustainability.[53] IOE will provide individual ratings for these IFAD-specific criteria. While Management will not rate "innovation", it will provide an individual rating to "scaling up", "environment and natural resource management" and "climate change adaptation" only for projects closing until IFAD12, to allow reporting on RMF11 and RMF12 targets agreed with Member States.

49 Apart from RIA evaluation products, all other evaluations use the aforementioned evaluation criteria and rating system.

50 Relevance, coherence, effectiveness, efficiency, impact and sustainability. See: https://www.oecd.org/dac/evaluation/daccriteriaforevaluatingdevelopmentassistance.htm

51 https://www.oecd.org/dac/evaluation/daccriteriaforevaluatingdevelopmentassistance.htm

52 In March 2022, IFAD adopted a new definition of innovation "IFAD defines innovation as a new process, product or approach that adds value and delivers a sustainable, equitable, inclusive and/or new contextual solution to rural development challenges. Furthermore, IFAD now defines its approach to innovation as follows: "IFAD aims to catalyse the generation, testing and scaling up of solutions that have the potential to contribute to deliver equitable, better and greater impact for the rural poor by leveraging on learning, strategic partnerships, digitalization and the implementation of suitable tools and guidelines. For IFAD, the most important innovations are those that impact rural poor people directly."

53 Knowledge management, partnership building and policy dialogue (assessed under coherence), innovation (assessed under effectiveness), scaling up and "environment and natural resources management and climate change adaptation" (assessed under sustainability), will continue to be rated individually. Part 2 of the manual provides guidance for assigning overall ratings at the project or country programme level.

78. Although Management will not rate these criteria using the six-point rating scale, it will monitor and measure them in COSOPs and at project completion and implementation. Management is redefining the approaches to these issues in IFAD12: the IFAD12 matrix of commitments includes an action plan on sustainability and a scaling-up strategy to be completed in 2022. The approach to natural resource management and climate change is captured in the revised Social, Environmental and Climate Assessment Procedures (SECAP) guidelines (2021, with roll-out in 2022) and in the relevant COIs and mandatory CIs in appendix – annex III. Turning to innovation, IFAD12 monitorable action 26 aims at developing an operating model and guidelines for innovation, to be overseen by the Change, Delivery and Innovation Unit; this will also be finalized in 2022. These 2022 actions are aimed at further refining the measurement approaches. Eventually, and in consultation with IOE, these objective measurements can serve as the foundation for Management "rating" these criteria.

79. Fourth, Management and IOE will follow different courses of action with regard to impact. IFAD Management assesses impact using a rigorous approach in five domains, in line with those included in the Tier II development indicators of its Results Management Framework (RMF):

i. Economic mobility;

ii. Strategic objective (SO)1: productive capacities (agricultural/non-agricultural production and productivity);

iii. SO2: access to market (access and integration into markets);

iv. SO3: resilience (ability of households to cope with climate and non-climate shocks); and

v. Improved nutrition.

80. As mentioned, Management measures impact through rigorous impact assessments conducted on 15 per cent of the portfolio that rely on a counterfactual-based analysis and guarantee attribution of results, thus representing a robust measure of the impact of IFAD-supported projects. In addition, Management reports the results measured through the COIs in the project logframe, which are also obtained through a rigorous survey methodology to determine attribution. As of IFAD12, all projects will conduct COI surveys at the baseline, midterm and completion stages (with treatment and comparison groups), since it is the best way to capture the results of a project's interventions over the course of its implementation. The core outcome indicators measurement guidelines (2021) provide the COI survey methodology. Table A in annex III presents the full list of COIs, mapped to the respective impact domain. Through the RMF, Management reports on the impact of its projects at the aggregate level. Since it presents quantitative evidence on the attributable impact of projects based on rigorous methodologies, as of IFAD12, Management does not rate impact.

81. IOE will continue to rate impact following international practice. IOE will draw evidence from impact studies conducted by projects, IFAD Management or other organizations and validate the findings independently, based on available information, fieldwork and its own expertise. In selected cases, IOE may conduct its own impact surveys. Similarly, IOE will take the data available through the COI surveys into account, validate them as above and collect additional data and information as required by the specific operations, development context or independent evaluation questions.

FIGURE 11

IFAD's evaluation criteria

Project-level evaluations	Country-level evaluations
International criteria Relevance Effectiveness • Innovation*	Criteria used for project-level evaluations and: the international criterion of coherence • Non-lending activities ▸ Knowledge management ▸ Partnership development ▸ Policy engagement
Efficiency Impact * • Changes in: incomes and assets; social/human capital; household food security and nutrition; institution and policies	
Sustainability • Scaling up* • Natural resources and climate change adaptation*	
IFAD-specific Gender equality and women's empowerment	
Partner performance • IFAD • Government	

* These criteria will continue to be rated by IOE and not by Management. With regard to scaling up and natural resource management and climate change, Management will rate these criteria only for projects closing until the IFAD12 period (2022–2024), in order to comply with RMF11 and RMF12 reporting requirements.

82. Table 1 presents the definitions of the criteria and related overarching evaluation questions. The use of core questions helps to ensure consistency and comparability. It allows for aggregation of ratings across IFAD evaluations and helps focus data collection and analysis.

TABLE 1

Definition of IFAD evaluation criteria

Evaluation criteria (project and country levels)	Overarching questions
RELEVANCE **The extent to which: (i) the objectives of the intervention/strategy are consistent with beneficiaries' requirements, country needs, institutional priorities and partner and donor policies; (ii) the design of the interventions/strategy* and the targeting strategies adopted are consistent with the objectives; and (iii) the intervention/strategy has been (re-)adapted to address changes in the context.** ***Evaluations will analyse the strategy pursued, whether explicit (written) or implicit.**	Was the intervention/country strategy and programme relevant and aligned with: (a) the country's development needs and challenges, as well as national policies and strategies; (b) IFAD's relevant strategies and priorities; (c) the needs of the beneficiaries and tailored to very poor or marginalized people or special categories. Was the design quality consistent with available knowledge, recognized standards (if available)? Was the design realistic in terms of suitability to the context and implementation capacity? Was the design re-adapted to changes in the context (if applicable)?
COHERENCE **(mainly for country level and strategic evaluations)** **This comprises the notions of external and internal coherence. External coherence is the consistency of the strategy with other actors' interventions in the same context. Internal coherence looks at the internal logic of the strategy, including the complementarity of lending and non-lending objectives within the country programme.** **Non-lending activities are specific domains for assessing coherence.** **KNOWLEDGE MANAGEMENT** **The extent to which the IFAD-funded country programme is capturing, creating, distilling, sharing and using knowledge.** **PARTNERSHIP BUILDING** **The extent to which IFAD is building timely, effective and sustainable partnerships with government institutions, international organizations, the private sector, organizations representing marginalized groups and other development partners to cooperate, avoid duplication of efforts and leverage the scaling up of recognized good practices and innovations in support of smallholder agriculture and rural development.** **POLICY ENGAGEMENT** **The extent to which IFAD and its country-level stakeholders engage, and the progress made, to support dialogue on policy priorities or the design, implementation and assessment of formal institutions, policies and programmes that shape the economic opportunities for large numbers of rural people to move out of poverty.**	How coherent is the country programme? In particular: To what extent were there synergies and linkages between different elements of the country strategy/programme (i.e. projects, non-lending activities)? How did IFAD position itself and its work in partnership with other development partners? Did IFAD contribute to policy discussion, drawing from its programme experience? To what extent have lessons and knowledge been gathered, documented and disseminated?
EFFECTIVENESS **The extent to which the intervention/country strategy achieved, or is expected to achieve, its objectives and results at the time of the evaluation, including any differential results across groups** **A specific subdomain of effectiveness relates to:** **Innovation, the extent to which interventions yielded a solution (practice, approach/method, process, product or rule) that is novel with respect to the specific context, timeframe and stakeholders (intended users of the solution), with the purpose of improving performance and/or addressing challenge(s) related to rural poverty reduction.[54]**	Were the objectives of the intervention/country strategy and programme achieved, or likely to be achieved, at the time of the evaluation? Did the intervention/strategy achieve other objectives or have any unexpected consequence? To what extent did the programme or project support/promote innovations aligned with stakeholders' needs or challenges? Were the innovations inclusive and accessible to a diversity of farmers (in terms of gender, youth and diversity of socio-economic groups)?

54 Conditions that mark an innovation: newness to the context, intended users and intended purpose of improving performance. Furthermore, the 2020 corporate-level evaluation on IFAD's support to innovations defined transformational innovations as "ones that are able to lift poor farmers above a threshold from which they cannot easily fall back after a shock." These innovations simultaneously tackle multiple challenges faced by smallholder farmers. In IFAD operational contexts, which happens by packaging/bundling several small innovations together. Most of the time, they are holistic solutions or approaches applied or implemented by IFAD-supported operations.

Evaluation criteria (project and country levels)	Overarching questions
EFFICIENCY **The extent to which the intervention or strategy delivers, or is likely to deliver, results in an economic and timely manner** **"Economic" is the conversion of inputs (e.g. funds, expertise, natural resources, time) into outputs, outcomes and impacts as cost-effectively as possible, compared to feasible alternatives in the context. "Timely" delivery is within the intended timeframe, or a timeframe reasonably adjusted to the demands of the evolving context. This may include assessing operational efficiency (how well the intervention was managed).**	Having considered the nature of the intervention and implementation context, key questions include (but are not limited to): What is the relationship between benefits and costs (e.g. net present value, internal rate of return)? How does it compare with similar interventions (if the comparison is plausible)?[55] Are the unit costs of specific interventions consistent with recognized practices and the results achieved? Are programme management cost ratios justifiable in terms of intervention objectives and results achieved, considering contextual aspects and unforeseeable events? Is the timeframe of the intervention design and implementation justifiable, considering the results achieved, the specific context and unforeseeable events?
IMPACT **The extent to which an intervention/country strategy has generated, or is expected to generate, significant positive or negative, intended or unintended, higher-level effects.** **The criterion includes the following domains:** **Changes in income, assets and productive capacity** **Changes in social/human capital** **Changes in household food security and nutrition** **Changes in institutions and policies** **The impact assessment will seek to determine whether changes have been transformational, generating changes that can shift societies onto fundamentally different development pathways (e.g. due to the size or distributional effects of changes to poor and marginalized groups).**	Has the intervention/country strategy and programme had the anticipated impact on the target group, institutions and policies? Why? What are the observed changes in target group incomes and assets, household food security and nutrition, social/human capital and institutions and policies over the project/COSOP period? What explains those changes? What are the challenges? From an equity standpoint, have very poor/marginalized groups and/or special categories, substantially benefited?
SUSTAINABILITY **The extent to which the net benefits of the intervention or strategy continue and are scaled up (or are likely to continue and be scaled up) by government authorities, donor organizations, the private sector and other agencies.** **Note: This entails an examination of the financial, economic, social, environmental and institutional capacity of the systems needed to sustain net benefits over time. It involves analyses of resilience, risks and potential trade-offs.** **Specific domain of sustainability:** **Environment and natural resource management and climate change adaptation. The extent to which the development interventions/strategy contribute to the enhancement of environmental sustainability and resilience to climate change in small-scale agriculture.** **Scaling up* takes place when: (i) bi- and multilateral partners, private sector and communities adopt and disseminate the solution tested by IFAD; (ii) other stakeholders invest resources to bring the solution at scale; and (iii) the government applies a policy framework to generalize the solution tested by IFAD (from practice to policy).** ***Note that scaling up does not relate to innovations only.**	To what extent did the intervention/country strategy and programme contribute to long-term institutional, environmental and social sustainability? What is the level of engagement, participation and ownership of the government, local communities, grass-roots organizations and poor rural people? In particular, did the government ensure budget allocations to cover operations and maintenance? Did the programme include an exit strategy?[56] For environment and natural resource management and climate change adaptation, to what extent is the intervention/strategy: • Improving farming practices? Minimizing damage and introducing offsets to counter the damage caused by those farming practices? • Supporting agricultural productivity that is sustainable and integrated into ecosystems? • Channelling climate and environmental finance through the intervention/country programme to smallholder farmers, helping them to reduce poverty, enhance biodiversity, increase yields and lower greenhouse gas emissions? • Building climate resilience by managing competing land-use systems while reducing poverty, enhancing biodiversity, increasing yields and lowering greenhouse gas emissions? **For scaling up:** To what extent were results scaled up or likely to be scaled up in the future? Is there an indication of commitment by the government and key stakeholders to scale up interventions and approaches – for example, in terms of the allocation of funds for selected activities, human resources availability, continuity of pro-poor policies, participatory development approaches and institutional support?

55 References to Management documents related to this criterion include: (i) the IFAD action plan for efficiency (ii) the IFAD Internal Guidelines on Economic and Financial Analysis of Rural Investment Projects, 2015; and (iii) IFAD's project implementation guidelines, appendix VII - value for money in supervision.

56 Useful references to Management's documents related to this criterion include the IFAD action plan on sustainability and the IFAD Project Design Guidelines, 2020 (notably appendix V).

Evaluation criteria (project and country levels)	Overarching questions
GENDER EQUALITY AND WOMEN'S EMPOWERMENT **The extent to which IFAD interventions have contributed to greater gender equality and women's empowerment. For example, in terms of women's access to and ownership of assets, resources and services; participation in decision-making; workload balance and impact on women's income, nutrition and livelihoods; and in promoting sustainable, inclusive and far-reaching changes in the social norms, attitudes, behaviours and beliefs underpinning gender inequality.** **Evaluations will assess the extent to which interventions and strategies have been gender-transformational relative to the context, by: (i) addressing root causes of gender inequality and discrimination; (ii) acting upon gender roles, norms and power relations; (iii) promoting broader processes of social change (beyond the immediate intervention).** **Evaluators will consider differential impacts by gender and the way they interact with other forms of discrimination (such as age, race, ethnicity, social status and disability), also known as gender intersectionality.[57]**	What were the project's achievements in terms of promoting gender equality and women's empowerment, including intersectionality issues? In particular, were there changes in: (i) women's access to resources, sources of income, assets (including land) and services; (ii) women's influence in decision-making within the household and community; (iii) workload distribution (including domestic chores); (iv) women's health, skills, nutrition? Were there notable changes in social norms, attitudes, behaviours and beliefs and policies/laws relating to gender equality? Was attention given to programme implementation resources and disaggregated monitoring with respect to gender equality and women's empowerment goals?
Partner performance (assessed separately for IFAD and the government) **The extent to which IFAD and the government (including central and local authorities and executing agencies) supported design, implementation and the achievement of results, a conducive policy environment and impact and the sustainability of the intervention/country programme** **The adequacy of the borrower's assumption of ownership and responsibility during all project phases, including government and implementing agency, for ensuring quality preparation and implementation, compliance with covenants and agreements, support for a conducive policy environment and for laying the foundation for sustainability and fostering participation by the project's stakeholders.**	**IFAD performance** How effectively did IFAD support the overall quality of design, including aspects related to project approach, compliance and operational aspects?[58] How proactively did IFAD identify and address threats to the achievement of project development objectives?[59] How effectively did IFAD support the executing agency in project management, financial management and setting-up project-level M&E systems? How did IFAD position itself and its work in partnership with other development partners? **Government performance** How tangible was the government's commitment to achieving development objectives and ownership of the strategy/project? Did the government adequately involve and consult beneficiaries/stakeholders at design and during implementation? How did the government position itself and its work in partnership with other development partners? How well did the project management unit (PMU)/project coordination unit manage start-up procedures, implementation arrangements, the appointment of key staff and resource allocation/funding? In how timely a manner did the PMU identify and resolve implementation issues? Was project management responsive to changes in the environment or the recommendations made during supervision missions or by the project steering committee? How useful were the various project management tools, annual workplan and budget (AWPB) and the management information system, developed during implementation? Were these tools properly used by project management? How did the PMU fulfil fiduciary responsibilities? How useful was the procurement plan and how was it used during implementation? How adequate were M&E arrangements made by the PMU, including the M&E plan, and the utilization of evaluation M&E data in decision-making and resource allocation?

57 Evaluation Cooperation Group, "Gender. Main messages and findings from the ECG Gender practitioners' workshops) (Washington, D.C., 2017), https://www.ecgnet.org/document/main-messages-and-findings-ieg-gender-practitioners-workshop

58 A useful reference to Management's comments that relate to the criterion is the IFAD project design guidelines (2020).

59 Sources for self-evaluations include Project Supervisions and Project Status Reports.

Ratings

83. Evaluation criteria are scored using a rating system introduced by IFAD in 2002. In 2005, IFAD moved from a four-point to a six-point rating system in line with the practice adopted in many other IFIs and United Nations organizations,[60] allowing for a more nuanced assessment of project results. In addition to performance reporting based on the six-point rating scale, in 2007 IFAD introduced the broad categories of "satisfactory" and "unsatisfactory" for performance reporting across the various evaluation criteria (see table 2).[61]

84. As a general rule, evaluators assign ratings, supporting their arguments with evidence and justifying the ratings with sound analysis. Evaluators are often faced with what is known as the "hindsight issue." This refers to the challenge of assessing (and rating) past projects with current perspectives, notably when applying evaluation criteria. Box 10 presents a way forward for dealing with the hindsight issue.

60 Evaluation Cooperation Group, *ECG Good Practice Standards for Evaluation of MDB Supported Public Sector Operations*. For each rated criterion, multilateral development banks use an even number (mostly four, exceptionally six, for greater differentiation) of rating scale points. For the sake of validity, credibility, transparency and comparability, they apply a clearly defined rating for each scale point that reflects a pre-defined set of ranked value terms.

61 The Annual Review of Development Effectiveness produced by the Independent Evaluation Group of the World Bank uses a similar categorization system.

BOX 10

Systems mapping

Evaluators are often faced with the challenge of assessing (and rating) past projects with current perspectives, especially when applying evaluation criteria. This is particularly challenging when the political or operational context, the operation targets and/or strategies have changed during the course of implementation. Holding managers accountable for failing to achieve today's standards before they were established may be unfair. For example, in cases where the context or policies changed late in the life of a project without the opportunity for course correction, prima facie it may seem anachronistic to assess managerial performance with today's metrics.

At the same time, it is well known that the context in which projects and country programmes are implemented can change. Change in complex systems is characterized by uncertainty, volatility and adaptation. In order to perform, a project or strategy must be able to adapt, not rigidly stick to its original formulation, when the context changes or is no longer conducive. Therefore, in several cases, it may be possible for programme managers to adapt to the context. In such cases, it is legitimate for an evaluation to assess the extent to which a programme has been resilient to the change in context or has been adapted to respond to changes.

All this requires a balancing act in an evaluation. On the one hand, it is fair to expect some capacity to adapt to changes in context. On the other hand, evaluators need to acknowledge when standards have substantially changed in recent times and cannot be applied retroactively.

Dealing with changes in contexts and standards over time (the hindsight issue)

85. Ratings for individual criteria are given using integers (i.e. no decimal points).[62] Consistent with most other evaluation offices and to keep the system simple, no weights are assigned to ratings when determining a final rating for overall project achievement. This manual provides the following general guidance (table 2) to support evaluators in the assignment of ratings on each evaluation criteria. With the aim of further reducing the space for subjectivity in ratings, as well as the disconnect between self- and independent evaluation, part 2 of this manual will provide further guidance on ratings.

62 Evaluations may establish composite ratings (e.g. arithmetic averages of other ratings) which would be rational numbers, with decimals. As an example, a rating for the overall project achievements could be established as the arithmetic average of all the individual project ratings.

TABLE 2

IFAD rating system

Score	Assessment	Indicative description	Category
6	**Highly satisfactory**	The activity, project, programme or policy achieved the vast majority of the main targets, objectives, expectations and results (or impacts). Due to its high quality, it could be considered an "outstanding practice" case. A rating of "6" may also signal that results (e.g. effectiveness, impacts) are "transformational."	
5	**Satisfactory**	The activity, project, programme or policy achieved the vast majority of the main targets, objectives, expectations and results (or impacts). In qualitative terms, the intervention/policy in question could be considered an example of good practice that inspires other programmes/policies.	**Satisfactory**
4	**Moderately satisfactory**	The activity achieved a relative majority of the targets, objectives, expectations, results or impacts. At the same time, there were some noticeable gaps in achievement. The quality of what was achieved was good, although not a special case of good practice.	
3	**Moderately unsatisfactory**	Under the criterion in question, the activity did not achieve a relative majority of its targets and objectives, results (or impact). There were areas of clear underachievement. In qualitative terms, the achievements were below standards and expectations.	
2	**Unsatisfactory**	Under the criterion in question, the activity achieved only a minority of its targets, objectives, expectations, results or impacts. The quality of achievement was low and well below standards.	**Unsatisfactory**
1	**Highly unsatisfactory**	Under the criterion in question, the activity (project, programme, non-lending) achieved almost none of its targets, objectives, expectations, results or impacts. The quality was very poor, and there may have been cases in which the situation was worsened.	

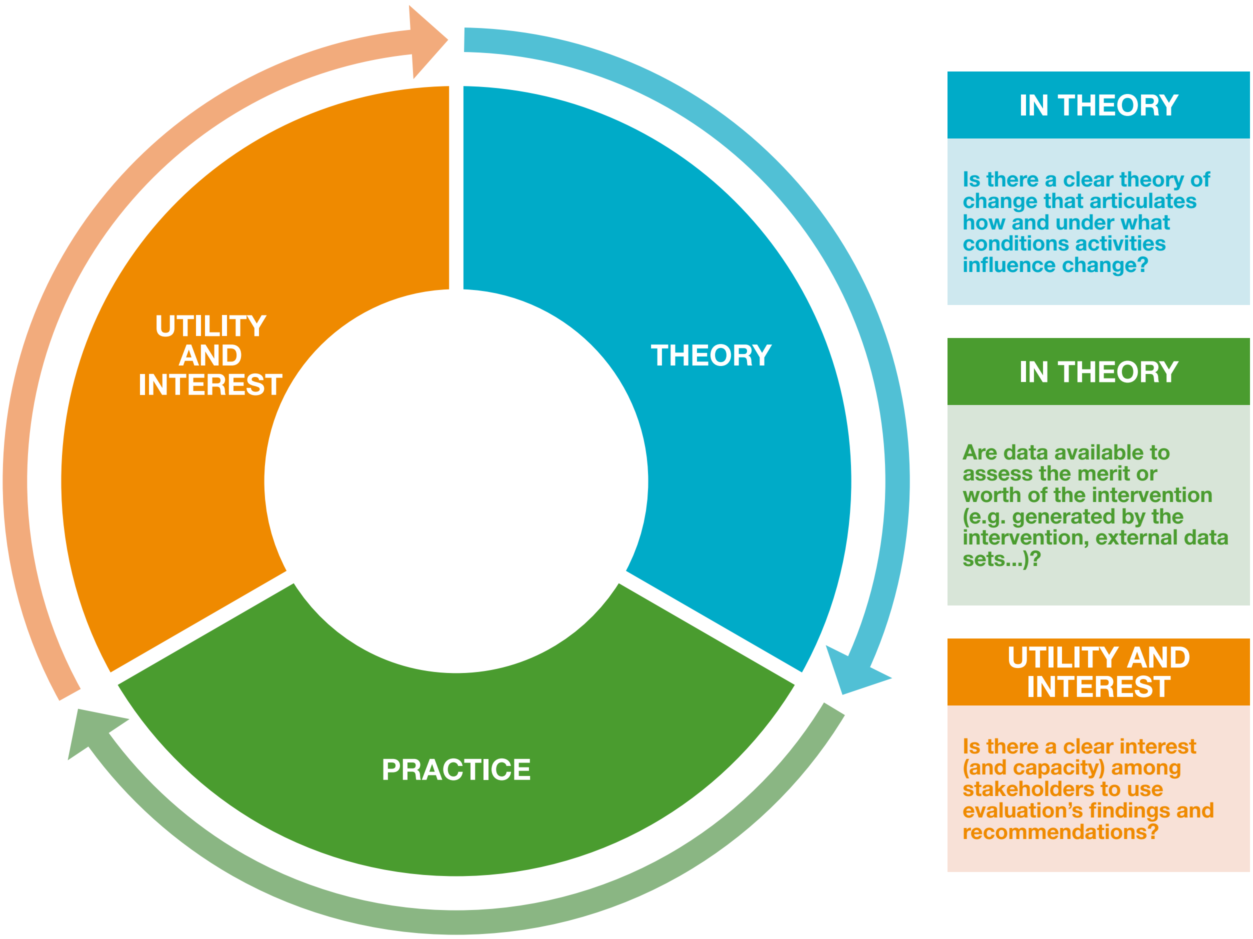
D. Reviewing evaluability and data availability

86. In the evaluation literature, the notion of evaluability assessment is related to an analysis to be conducted before deciding whether a specific evaluation should be undertaken and when.[63] A graphic example is displayed in figure 12.

63 OECD-DAC, Glossary of key terms in evaluation and results-based management (Paris: OECD-DAC, 2010), 21.

FIGURE 12

What do evaluability assessments examine?



Source: Adapted from Betterevaluation.org.

87. However, the reality in many development organizations is that the decision to conduct a specific evaluation is often driven by governing bodies and corporate-level commitments. Thus, the assessment of what can be evaluated often takes place after an evaluation has been approved. As discussed at the beginning of this chapter, the examination of the scope, approach and evaluability are in part overlapping and iterative processes. Thus, it is important to recognize that the review of evaluability may have a feedback loop on the scope of an evaluation.

88. For many evaluations, the review of evaluability and data availability typically focuses on:

i. Whether it is timely for assessing the results of a policy, strategy or project, or whether the evaluation should be confined to the ongoing process and the likely pathway toward achieving certain results;

ii. Whether secondary data are available to complete the analysis according to specific evaluation criteria and what are the main information gaps to be filled in and through what methods; and

iii. To what extent it will be possible to collect and analyse certain data, depending on time, budget and other circumstances (e.g. security, credibility, social acceptance).

89. Thus, the assessment of evaluability and data availability is an important consideration when making decisions on data collection and analysis. Relevant inputs include the review of background documentation and databases (e.g. World Bank, United Nations system, think tanks, literature), official documentation from IFAD, government and other agencies, surveys and preliminary interviews with the main stakeholders (both in person and virtual). In some cases, a brief reconnaissance mission to a country or project site may be required to complete the review.

E. Approaches and methods for data collection and analysis

90. Data collection tools vary with the type of evaluation scope, approach, outcome of the evaluability assessment, availability of secondary data and other contextual factors.

91. Mixed methods for analytical rigour and depth. In line with international good practice,[64] IFAD encourages triangulation of methods, data collection and data analysis. All evaluations must be evidence-based and explicitly consider limitations related to the analysis conducted. Evaluators will always strive to identify and use the methods best suited to the specific purposes and context of the evaluation and consider how other methods may compensate for any limitations of the methods selected. In particular, strategic and aggregate-level evaluations – e.g. multilevel, multisite evaluations in country, thematic and cluster evaluations – are by definition multimethod. However, the idea of informed evaluation design, or the strategic mixing of methods applies to all evaluations.

92. Using a mix of designs and methods and triangulating information from different approaches is recommended to assess different facets of complex outcomes or impacts, as well as to capture the cultural and contextual complexities that affect achievement of the desired goals.[65,66] This yields greater validity than a single method.[67]

64 Evaluation Cooperation Group, ECG Big Book on Good Practice Standards (Washington, D.C., 2012).

65 C. Bolinson, D.M. Mertens, *Transformative evaluation and impact investing: A fruitful marriage.* In R.P. Herman & E. de Morais Sarmento (Eds.).Global Handbook of Impact Investing (Wiley, 2021).

66 Hur Hassnain, Lauren Kelly and Simona Somma, eds., *Evaluation in Contexts of Fragility, Conflict and Violence: Guidance from Global Evaluation Practitioners* (Exeter, UK: IDEAS, 2021).

67 For example, UN Women evaluated its contribution to gender-responsive budgeting in Europe and Central Asia by combining a participatory ToC approach with outcome harvesting – an approach for capturing unintended positive and negative outcome-level results of interventions within complex contexts. To strengthen the depth and utilization of the analysis, the evaluation also attempted to calculate a social return on investment for gender-responsive budgeting (GRB). This evaluation also enriched the analysis by applying, for example, stakeholder mapping using a sequenced application of critical systems heuristics, human rights role analysis and a power analysis; mapping the history of gender-responsive budgeting in the region, including a rich picture (systems view) of the forces at play in reaching the decisions made; and an institutional pathways analysis assessing the system dynamics that have influenced the history of GRB in the region and frame the options for future change. See, for example: UNWOMEN (2017) Evaluation of UNWOMEN's contribution to gender-responsive budgeting in the Europe and Central Asia Region.

93. Keeping cultural responsiveness at the forefront of evaluation efforts is key to broadening participation and incorporating culture and context into an evaluation. This requires evaluators to be sensitive and responsive to the cultural context in which the programme and/or policy is operationalized at all stages of the evaluation process. This is particularly important in instances when the participants' culture is known to have a major influence on outcomes. Ensuring systematic and coherent application of a culturally responsive evaluation begins with integrating cultural dimensions into the evaluation framework (box 11).[68] Since culturally responsive evaluation is an emerging approach, UNEG has identified key questions that could be posed as a minimum, and aspirational questions that could be considered when designing and implementing an evaluation.[69]

68 M. Bryan, A. Lewis, A. Culturally Responsive Evaluation as a Form of Critical Qualitative Inquiry. *Oxford Research Encyclopedia of Education* (2019).

69 For a list of culturally responsive evaluation questions and approaches see: UNEG Development of culturally responsive criteria for evaluations (2018), **http://www.unevaluation.org/document/detail/2123**

BOX 11

Systems mapping

Culturally responsive evaluations are based on the notion that evaluation cannot be separated from the sociocultural contexts within which programmes are implemented. Culture shapes the behaviours and worldviews of its members and is therefore central to our understanding of individuals' motivations, attitudes and responses to an intervention […] To be culture-blind in evaluation runs the risk of perpetuating inequalities, in the same way that gender-blind evaluation or policy does.

Source: UNEG (2018) Development of culturally responsive criteria for evaluations.

94. For evaluations to adopt participatory[70] and culturally responsive approaches,[71] understanding the context and engaging stakeholders are key steps for preparing the entire evaluation process, including data collection, analysis and the dissemination of results (figure 13). The most appropriate way of approaching primary stakeholders will largely depend on the local dynamics, socio-economic settings and customs. Evaluators need to understand the context in which the evaluation activities take place and adapt accordingly, placing culture and the community at the centre of the evaluation.

70 For examples and documents that discuss good practices for participatory methods in evaluation see **https://www.participatorymethods.org/**

71 For more information and examples of culturally responsive evaluation, see: J.A Chouinard and F. Cram *Culturally Responsive Approaches to Evaluation* (USA: Sage Publications, 2020); D. Mertens, *Mixed Methods Design in Evaluation* (USA: Sage Publications, 2018); UNEG, *Compendium of Evaluation Methods Reviewed Volume 1* (New York, 2020); Frey, B. Culturally Responsive Evaluation (2018); and the Sage Encyclopedia of Educational Research, Measurement, and Evaluation, Vol. 4. A recent application can also be found in B. Chilisa and D. Mertens, "Indigenous Made in Africa Evaluation Frameworks: Addressing Epistemic Violence and Contributing to Social Transformation," *The American Journal of Evaluation*, Vol. 42(2) 241-253 (2021).

FIGURE 13

Culturally Responsive Evaluation Framework

Source: Adapted from Frierson et al. (2002) and Hood et al. (2015), quoted in Chouinard and Cram (2018).

95. **An integrated gender focus is also fundamental to the collection of relevant information**. In addition to strengthening validity through the triangulation of different data collection methods, mixed methods are particularly important for gender- and equity-responsive evaluation to: (i) study empowerment and behavioural change processes that are hard to capture with a single data collection method; (ii) strengthen the generalizability of in-depth qualitative analysis (i.e. to ensure the sample of respondents is representative of the total sample population).[72]

96. Adopting participatory methods and data collection tools for evaluation and systematically developing evaluation frameworks that include the voice of marginalized people are a key component of IFAD's evaluations. This not only helps to accurately collect the voice of underrepresented groups but increases the validity and reliability of the evaluation. Participation can occur at any stage of the evaluation process: in its design, data collection, analysis and reporting; it is not exclusive to specific evaluation methods or restricted to quantitative or qualitative data collection and analysis.[73] The needs and decisions about the type and extent of participation are usually different for an evaluation that focuses on local-level impacts and one that examines national-level change. IFAD evaluations need to pay attention to promoting the participation of key stakeholders in the evaluation process, but at the same time ensure that the principles of impartiality, credibility and transparency of the evaluation's analysis and final judgments are upheld.

72 Evaluation Cooperation Group, *Integrating gender into project-level evaluation*. ECG reference document. (Washington, D.C., 2017).

73 I. Guijt, Participatory Approaches, *Methodological Briefs: Impact Evaluation 5* (Florence: UNICEF Office of Research, 2014).

97. Figure 14 presents commonly used data collection methods. The list is not exhaustive, and a specific evaluation product might need a particular data collection method that is not included in the list. Most evaluations use a combination of methods, as needed.

98. In the context of strategic and aggregate-level evaluations, designs may cover different case-study levels, with cross-case (comparative) analysis across countries (or interventions). Case studies are often used in IFAD evaluation, although they present a number of methodological challenges, particularly in terms of internal and external validity. With regard to internal validity, the concern is how to ensure the quality, reliability and robustness of methods and design. The concern with regard to external validity is generalizations – i.e. the extent to which it is possible to generalize, and under what circumstances. Lastly, there are issues related to aggregation and synthesizing for learning purposes. There is an ongoing international debate on this topic and different ways to analyse and synthesize findings. One such approach is qualitative comparative analysis (see box 12).[74]

74 V. Pattyn, A. Molenvend, B. Befani, "Qualitative Comparative Analysis as an Evaluation Tool: Lessons from an Application in Development Cooperation," *American Journal of Evaluation*, vol. 40(l) 55-74 (2019).

BOX 12

Qualitative comparative analysis (QCA)[75]

Primarily designed to answer the questions: (i) Under what circumstances did the programme generate or fail to generate the desired outcome? and/or (ii) What works best, why and under what circumstances? In essence, QCA is a case-based method that enables evaluators to systematically compare cases that are responsible for the success or failure of an intervention by identifying key factors in each case. What differentiates this approach from most other cross-case comparative methods is that it provides a specific set of algorithms to analyse data sets (usually in the form of a table) by using Boolean algebra logic operators to document varying configurations of conditions associated with observed outcomes. In this sense, QCA can also be considered a data analysis technique.

QCA is usually designed for use with an intermediate number of cases – typically between 10 and 50. It is not appropriate in all circumstances, as it requires a strong ToC and clearly defined cases and cannot measure the net effects of an intervention.

Example:

- Impact Evaluation of the Global Environment Facility Support to Protected Areas and Protected Area Systems (September 2016). Here, the evaluators used a theory-based design combining multiple methods, including multilevel analysis (global and portfolio) and qualitative comparative analysis. Available here.

Source: Valérie Pattyn, Astrid Molenveld, Barbara Befani, Qualitative Comparative Analysis as an Evaluation Tool: Lessons from an Application in Development Cooperation (2017).

75 Source: Valérie Pattyn, Astrid Molenveld, Barbara Befani, Qualitative Comparative Analysis as an Evaluation Tool: Lessons from an Application in Development Cooperation (2017).

FIGURE 14

Frequently used data collection methods

		WHEN TO USE
	DOCUMENT REVIEW	To identify available data by reviewing formal policy documents, M&E reports, programme records, political, socio-economic agricultural profiles of the country or specific locale.
	DIRECT OBSERVATION	To learn how the programme naturally occurs by observing sites, practices, living conditions, physical constructions using a well-design observation record form (notes, photos or video).
	PHYSICAL MEASUREMENT	To measure physical changes based on agreed indicators and measurement procedures. Examples include birth weight, nutrition levels, rain levels, and soil fertility.
	INTERVIEWS	To understand individual experiences in more detail. Can be unstructured, semi-structured or structure questions.
	SURVEY	To collect information from a defined group. They are standardized instruments and are usually comprised of well-defined, close ended questions. Can be administered in person, mail, telephone.
	FOCUS GROUP DISCUSSION	To discover issues that are of most concern for a community or group.
	CASE STUDY	To examine in-depth a limited number of cases. Useful for documenting contextual conditions and producing insights about whether the programme might make a difference in other settings.
	MEMORY RECALL	To reconstruct beneficiaries and other stakeholders, situation before the project.
	SYSTEMATIC REVIEW	To gather all available empirical data by using clearly defined, systematic methods to obtain answers to specific question.
	EVIDENCE GAP MAPS	To identify key “gaps” where few or no evidence from impact evaluations and systematic reviews is available.

Note: This illustration provides a selection of available methods. For a rapid review of methods, see, for example, Vaessen, et al. (2020), IEG World Bank, op cit.

F. The role of information and communications technology for data collection and analysis

99. Evaluators across all regions of the world face recurring challenges in the field. Lack of reliable M&E data, limited time and resources and operating in contexts that are often fragile and affected by conflict and violence are some of the more common obstacles. The COVID-19 pandemic and related travel restrictions have adversely affected the design, implementation and evaluation of international development interventions and the ability of the evaluation function to capture the consequences of the economic crisis facing the rural poor and marginalized people.[76] New technologies for data collection and analysis (and new types of data) are slowly but steadily making their way into international development practice and its evaluation. This is an area of growing interest for IFAD. In 2017, IOE organized an international conference on ICT for evaluation and published a book on the subject.[77] In 2019, IFAD prepared a Strategy for Information and Communications Technologies for Development.[78]

100. The growing emphasis on complexity, real-time feedback and adaptive management approaches (see chapter II), coupled with the COVID-19 pandemic, have accelerated remote data collection to minimize the risk of spreading the virus and underscored the urgency of getting accurate data quickly. ICTs are offering new methods and tools for gathering, analysing and disseminating data and are changing the way evaluations are conducted, potentially opening the door to more rigorous evaluation.

101. Evaluators now have a variety of tools at their disposal that enable more data to be collected, often remotely, and processed faster. A comprehensive description of the vast array of emerging data collection technologies is beyond the scope of this manual, but links to other sources of information are provided, where relevant.[79] Figure 15 presents a summary of the most prominent tools and methods for data collection and analysis. Different tools have specific strengths and weaknesses. Typical opportunities offered by ICT-inspired innovations for evaluation involve data collection and data analysis. There are also opportunities for data and information display and communication activities.

76 https://www.ifad.org/documents/38714182/42217951/LearningNote_Covid19_forweb2.pdf/98f22bb0-6c22-16c3-c54b-4f09b4f0fdcd?t=1610977391000

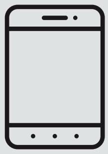
77 O. Garcia, P. Kotturi, eds., *Information and Communication Technologies for Development Evaluation* (Routledge, 2019).

78 https://webapps.ifad.org/members/eb/128/docs/EB-2019-128-R-5.pdf

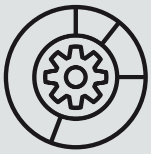
79 For a detailed discussion of the role of ICTs and big data in evaluation practice, see, for example: P. York, M. Bamberger, *Measuring results and impact in the age of big data: the nexus of evaluation, analytics, and digital technology* (New York: The Rockefeller Foundation, 2020); and Hassnain, H., Kelly, L., Somma, S., eds. "Evaluation in Contexts of Fragility, Conflict and Violence Guidance from Global Evaluation Practitioners" (IDEAS, 2021).

FIGURE 15

ICT tools: Advantages and disadvantages[80]

ICTs for data analysis			
	Description	**Advantages**	**Disadvantages**
MOBILE DATA COLLECTION 	The targeted gathering of structured information using mobile phones, tablets or personal digital assistants through a special software application.	It can improve the timeliness and accuracy of the data collection. Platforms allow customization of surveys to include photographs, voice recordings and GPS coordinates.	Technology alone will not improve the survey design or instrument. Potential bias in favour of well-educated or well-off citizens.
REMOTE SENSING 	Observing and analysing a distant target using the electromagnetic spectrum of satellites, aircraft or other airborne devices.	The ability to collect data on inaccessible areas. Observed objects or people are not disturbed.	Privacy concerns Potentially high cost of obtaining images or primary data collection using remote sensors. Socio-economic indicators hard to capture.
GEOGRAPHIC INFORMATION SYSTEMS 	Computer-based tool for integrating and analysing geographic or spatial data.	Combination of different types of geographical data sets. It allows viewing, interpreting and visualizing data a number of ways – revealing relationships, trends and patterns. GIS can also be used to digitally represent and interpret oral and life histories and can accommodate qualitative information.	GIS set-up is complex. In addition to the cost of the equipment, there is the training cost. Frequent updating of datasets or data models may lead to errors in results.
CLOUD COMPUTING 	Delivery of computing services – servers, storage, databases, networking, software, analytics and more – over the Internet ("the cloud"), thus permitting shared access to resources.	Access to data storage and analytical tools in a shared manner enables organizations to operate effortlessly across geographical areas. Real-time integration of data collection, analysis and reporting	Security concerns over access to data. Requires a robust high-speed internet connection.
DATA VISUALIZATION 	Representation of data graphically and interactively.	Graphic and interactive presentation of data increases the accessibility of complex data sets and, in turn, the use of the data. It can identify trends and patterns in complex and large data sets.	Identifying and putting together data visualization can be time-consuming or costly if outsourced.
MACHINE LEARNING 	A set of methods for getting computers to recognize words and images, and creating prediction models.	Reviews large volumes of data and identifies patterns, trends or specific information.	Needs large data sets to train on. Initial algorithm training is time consuming.

80 Multiple sources, see for instance: UNDP, *Discussion Paper: Innovations in Monitoring & Evaluating Results* (New York, 2013). INTRAC, ICT in Monitoring and Evaluation (Oxford, 2017).
O. Garcia, P. Kotturi, eds., *Information and Communication Technologies for Development Evaluation* (Routledge, 2019).

ICTs for data analysis			
	Description	**Advantages**	**Disadvantages**
BIG DATA ANALYTICS 	The use of advanced analytic techniques against very large, diverse big data sets that include structured, semi-structured and unstructured data from different sources of different sizes.	Access to a range of descriptive, exploratory and predictive analytics tools, which makes it possible to develop models, evaluate complex programmes and predict future trends.	"Data exhaust" is not representative of the wider population, much less of the marginalized.

102. ICTs offer an unprecedented number of options for evaluators to access, gather and analyse data more efficiently. ICTs enable evaluators to go further in exploring the ToC (see chapter III.B) and to do so with greater rigour. They are therefore critical to strengthening evidence-based policy making that relies on evaluation findings. Evaluators need to keep abreast of the available ICT tools in order to best decide when and how to incorporate them into their work.[81]

103. ICTs are not a panacea, however, but a means to an end. Technology is only as good as the evaluators who use it; and the evaluations of development programmes will still need to be grounded in robust theory. ICTs can also risk increasing biases, where assumptions included in the computing models can lead to a false sense of objectivity regarding the results. Evaluators must be aware of inherent biases that may be built into the data collection and coding processes and/or the software used to analyse the data.

104. Furthermore, the introduction of ICTs cannot be viewed as a stand-alone phenomenon but as a part of an organization-wide process and must be mainstreamed into IFAD's operations. This may include mainstreaming technology into planning, M&E and self-assessment processes.

105. Issues surrounding data privacy, ethics and inclusiveness related to the use of ICTs for evaluation will need to be taken into consideration. In IFAD, the following principles will drive the selection of ICT tools for evaluation:[82]

- People-centric: keeps the interests of target groups at the centre of ICT use. Evaluations will not use technology solely for the sake of innovation.

- Inclusive: serves to include vulnerable and marginalized populations in the evaluation process. The issue of power discrepancy between those who produce the data and those who use it is vital in this context.

- Mixed-methods: combines traditional, participatory face-to-face data gathering with technology-enabled data collection methods and larger-scale data analytics. This addresses concerns about inclusiveness, makes sense of what big data patterns are showing and what might be missing from big data sources and ensures that important contextual clues are not missed.

- Privacy and ethics: protecting privacy and following ethical guidelines on how information is collected and shared to ensure that beneficiaries are not put at risk.

81 O. Garcia, P. Kotturi, eds. *Information and Communication Technologies for Development Evaluation* (Routledge, 2019).

82 These principles draw from several sources, in particular, L. Raftree, L., *Technology, Biases and Ethics: Exploring the Soft Sides of Information and Communication Technologies for Evaluation* (ICT4Eval, 2019), in O. Garcia, P. Kotturi, (eds), op.cit.

G. Evaluation conclusions and recommendations

106. Each evaluation should clearly present conclusions in the form of key messages that are informed by the main findings but are not a repetition or simple summary of the findings. Conclusions bring findings in the report to a higher level. They add value to the findings by providing an answer to the overarching questions of the evaluation. They also provide an explanation for the findings, highlighting the main underlying factors.

107. Conclusions help to bridge the findings and recommendations. However, conclusions should be kept separate from the recommendations, both in content and language (e.g. they should not state what ought to be done to improve a certain situation).

108. Conclusions are more forceful when they concentrate on a limited number of judgment statements (indicatively, three to six) that take the overall findings of the report into consideration and point to the main lessons from the evaluation: what worked, what did not and what the key factors were. This helps the transition to the recommendations.

109. Recommendations are proposals for action made to entities in charge of a programme, a strategy and/or policies to bring about improvements in performance and results. The quality of recommendations is a critical factor in evaluation to optimally stimulate learning, accountability and organizational effectiveness. The UNEG's Improved Quality of Evaluation Recommendations Checklist (2018) defines evaluation recommendations as proposals aimed at enhancing the effectiveness, efficiency, impact, relevance, sustainability, coherence, added value or coverage of the operation, portfolio, strategy or policy under evaluation. Recommendations are intended to inform decision-making, including programme design and resources allocations.[83] To this end, developing recommendations involves weighing effective alternatives, policy and funding priorities within a broader context. It requires in-depth knowledge of the context, particularly the organizational context in which policy and programme decisions will be made and the political, social and economic context in which investments operate.

110. Care must be taken to ensure that recommendations are: (i) appropriate for achieving the objectives of the interventions; (ii) few in number (typically from three to six); (iii) positioned strategically; and (iv) once implemented, will add value to the organization. Recent guidance on preparing recommendations is available from the UNEG's Improved Quality of Evaluation Recommendations Checklist (2018) (box 13) and from ECG Practice Note Formulation of Evaluation Recommendations (2018).[84]

83 UNEG "Improved Quality of Evaluation Recommendations Checklist". *Working paper* (2018). This checklist includes useful background information, and it also deals with the follow-up of evaluation recommendations, using the UNEG *Good Practice Guidelines for Follow-up*.

84 Evaluation Cooperation Group, *ECG Practice Note Formulation of Evaluation Recommendations* (Washington, D.C., 2018).

BOX 13

UNEG quality standards for recommendations

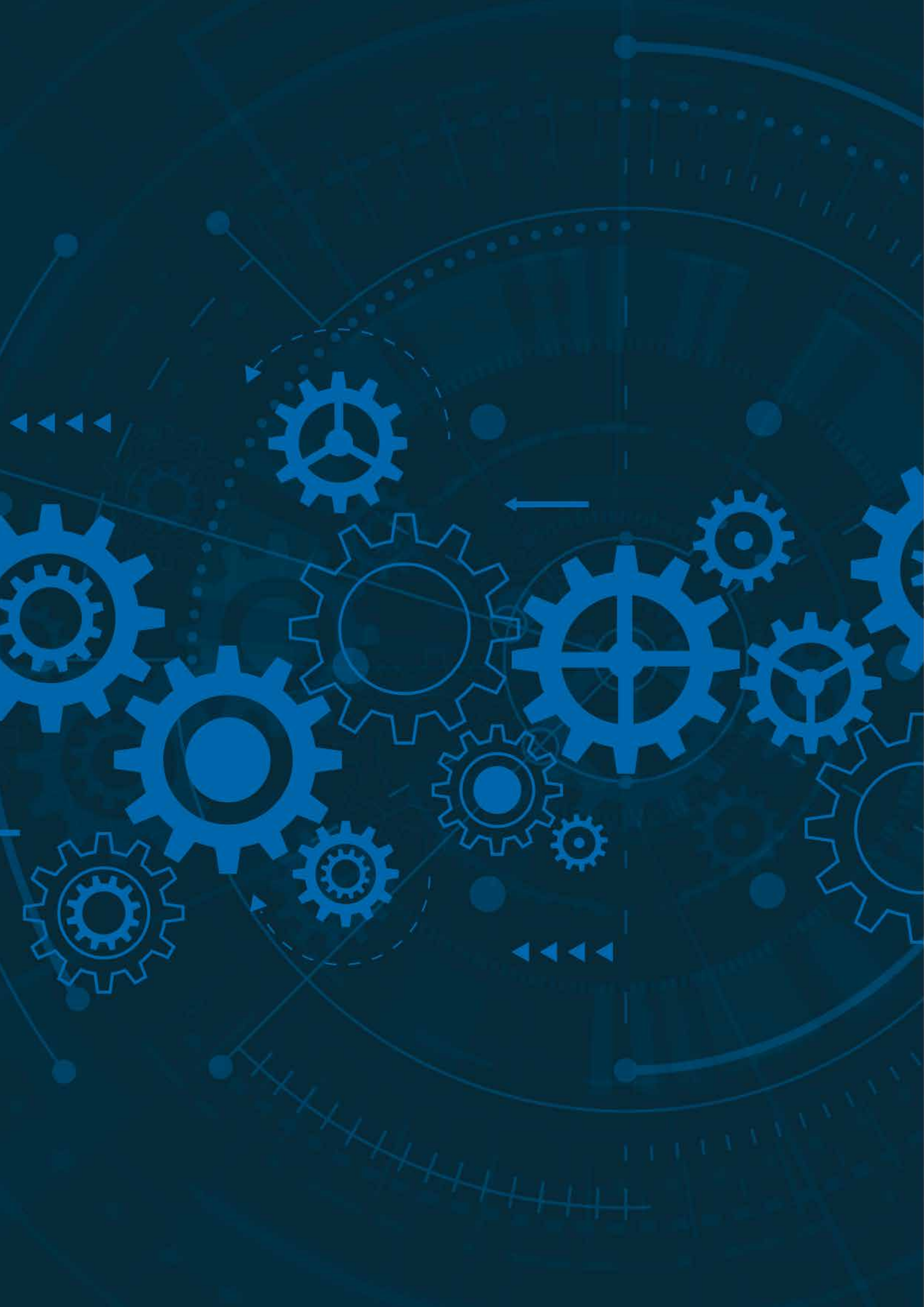
- The report describes the process followed in formulating the recommendations, including consultation with stakeholders.
- Recommendations are firmly based on evidence and conclusions.
- Recommendations are relevant to the objectives and purposes of the evaluation.
- Recommendations clearly identify the target group for each recommendation.
- Recommendations are clearly stated, with priorities for action made clear.
- Recommendations are actionable and reflect an understanding of the commissioning organization and potential constraints to follow up.

Source: UNEG Improved Quality of Evaluation Recommendation Checklist (2018).

111. As noted earlier in this manual, the full utility of an evaluation hinges on participation, dissemination, learning and follow-up. Recommendations must therefore be presented in a way that allows different decision-makers to clearly identify their responsibility. This should also facilitate tracking of follow-up actions by Management in the annual President's Report on the Implementation Status of Evaluation Recommendations and Management Actions (PRISMA). IFAD Management is also planning to set up an online system for tracking recommendations.

112. Part 1 (Chapters I-III) of this manual has provided the overall context for evaluation in IFAD, addressing the foundational elements of IFAD evaluation, including its mission, evaluation objectives, architecture, frameworks and the principles and criteria that guide all evaluations in IFAD. Part 2 provides practical and detailed guidance on different evaluation products covering both independent and self-evaluations, as well as the linkages between them.

Annexes



I. Annex
IFAD internal evaluation architecture

Self-evaluation system

1. At the core of self-evaluation is the Development Effectiveness Framework (DEF). Introduced in 2016, the DEF was developed to ensure that evidence is collected from projects and systematically used and to create the necessary structure to facilitate the collection and use of evidence in decisions on project design and implementation. Self-evaluation projects are designed to achieve the expected results of the DEF – namely, to strengthen accountability, enhance learning and ultimately, ensure that the decision-making process is based entirely on reliable evidence.[1]

2. To ensure relevance to the Twelfth Replenishment of IFAD's Resources (IFAD12) business model, which is centred on transformational country programmes and supported by institutional change and a revised financial framework, an updated version of the DEF will be introduced in 2022. The updated DEF provides the framework for improving IFAD's self-evaluation structure around three key pillars: (i) enhancing monitoring, evaluation, adaptation and learning; (ii) enhancing capacity, mainstreaming, sustainability, efficiency and scaling up; and (iii) working at the country level to maximize impact beyond projects. In line with the updated DEF and IFAD12 commitments, IFAD is developing dedicated action plans for areas where the Independent Office of Evaluation of IFAD (IOE) and Management have consistently found project/programme performance to be weak: sustainability, efficiency, scaling up, monitoring, evaluation, adaptation and learning. Updates will also be made in areas such as working under conditions of conflict and fragility and IFAD's offer on country-level policy engagement.

3. Self-evaluation products are developed at three main levels: country, project and corporate.

4. **At the country strategic opportunities programme (COSOP) level**, self-evaluation begins at design, when the results framework for the country strategy is reviewed by IFAD's Quality Assurance Group, Operational Policy and Results Division (OPR) and other members of the Operational Strategy and Policy Guidance Committee, utilizing the development effectiveness matrix for COSOPs. Light-touch reviews are conducted every year, and midway through implementation. COSOP results reviews are conducted to assess progress toward results, lessons learned, risk factors encountered and changes in country demand and priorities. At completion, COSOPs undergo a completion review – i.e. a self-evaluation of their strategic objectives and IFAD's performance in achieving them. Lessons learned from IFAD engagement fuel the preparation of new COSOPs.

5. **At the project level**, self-evaluation is fully integrated into the operation life cycle. At design, the development effectiveness matrix is used to review and enhance evaluability. To this end, the operation's theory of change and logical framework, including impact, outcome and output indicators, are reviewed, together with their baseline and target values.

1 https://webapps.ifad.org/members/ec/115/docs/EC-2021-115-W-P-6.pdf

6. During implementation, project teams prepare the annual supervision report, describing the progress made and identifying the main challenges encountered during execution. They also update progress data against logframe indicators and targets and rate project performance based on a set of pre-defined criteria. Following an adaptive management approach, such information is used at the project level to identify corrective actions and adjust the annual work plan and budget, including through the creation of a project improvement plan, if needed. At the portfolio-management level, the information from the logical frameworks and project supervision reports is used to ensure that adequate expertise and budget are allocated where performance requires follow-up or correction. At midterm, project teams conduct a full review of the progress made and report it in the midterm review; based on the evidence collected, the logframe is updated as needed and relevant decisions regarding the future of the operation are made, including possible restructuring.

7. At the end of the operation execution period, the relevant regional division prepares a project completion report (PCR). Through the PCR, project teams (under previous practice, government actors/the project management unit) rate the operation's development effectiveness according to the standard criteria and additional ones in line with IFAD strategies, including rural poverty impact, environment and natural resource management, climate change adaptation and gender equality. PCRs also include a section on lessons learned to benefit the design of new operations and improve implementation by building on experience.

8. In addition to these common self-evaluation practices, which are applied to all projects, the Research and Impact Assessment Division (RIA) conducts rigorous impact assessments of a representative sample of approximately 15 per cent of the projects closing in each replenishment period. Impact assessments use non-experimental methods to estimate the attributable impact of individual projects on IFAD strategic goals and objectives. Qualitative methods are also used to provide additional information on the context to complement the analysis. IFAD's impact assessment agenda is an important component of self-evaluation at the project and corporate levels. Furthermore, individual project impact assessments enrich the PCR evaluation of the rural impact.

9. Under the updated DEF, and in line with IFAD's graduation policy, COSOP guidelines are being updated and improved to foster adaptation and learning tools and enhance country-level sustainability, scalability, partnership and policy[2] influence. In addition, project supervision guidelines are being revised with the integration of tools for better data collection and monitoring, with special attention to the use of geographic information systems/satellite data to enhance monitoring and evaluation systems. Moreover, the existing guidelines to collect core outcome indicators[3] are increasingly being integrated into project design and will receive special emphasis in the revised supervision guidelines. PCR guidelines are also being updated to reflect a more objective scoring scale, as well as to benefit more from RIA impact assessments (in terms of data and estimated attributable impacts) in cases where the project is part of the impact assessment sample. The responsibility for preparing the PCR is being shifted from governments to IFAD, which is expected to improve PCR quality, candour, timeliness and transparency.

10. **At the corporate level**, the Report on IFAD's Development Effectiveness presents the Fund's annual operational and organizational performance by reporting on a set of 79 Results Management Framework (RMF) indicators agreed upon with Member States. IFAD also conducts thematic or cluster reviews on areas of specific interest to the Fund, using data from ongoing projects on a continuous basis through stocktaking to emphasize areas/countries/regions in which performance requires attention. This emphasis on real-time data contributes to IFAD's culture of results beyond specific self-evaluation products.

2 https://webapps.ifad.org/members/eb/133/docs/EB-2021-133-R-5.pdf

3 A set of indicators that measure the expected change resulting from beneficiaries' participation in the project, collected through a rigorous methodology that demonstrates attribution of results.

11. In addition, the attributable impacts estimated by RIA for a sample of projects are aggregated in a meta-analysis at the end of each replenishment period. This then feeds into a projection exercise to calculate the attributable impact of IFAD's overall portfolio during that period. The results contribute to corporate reporting and learning to improve future design and targeting for better impact.

12. Information on most self-evaluation products is monitored and captured through online systems, each one dedicated to a specific purpose and stage of the project life cycle, from concept to ex-post evaluation (see table 1). Data from self evaluation products are also presented in dedicated dashboards for both internal and external audiences.

TABLE 1

IFAD data systems

Operation Document Centre	Internal corporate system for sharing and managing operation documents produced in all IFAD interventions. It is designed to manage project and programme documents from design through completion, as well as documents across global/corporate, regional and country portfolios.
Grants and Investment Projects System (GRIPS)	Internal corporate system for keeping a record of projects financed through investment or grant programmes, together with their financial information.
Operational Results Management System (ORMS)	Internal system for the management and tracking of quantitative and qualitative project information related to: logframe indicators – baselines, targets and progress data; performance during implementation; development effectiveness at completion; action tracker; and lessons learned.
Quality Assurance Archiving System	Internal platform for managing the quality assurance review of all project designs and soon, of all grants, concept notes and COSOPs.
Commitment tracker	Internal tool used to track monitorable actions and outputs to fulfil commitments made under IFAD's replenishment. First formulated for IFAD11, it is being replicated for commitments agreed upon for IFAD12 (2022-2024).
Operations dashboard	Internal dashboard that provides up-to-date information on the performance of IFAD-supported country strategies and projects/programmes. Displays data on design, implementation and performance rating (with most external data published on IFAD website).
RMF dashboard	Section on the IFAD website where all donors, stakeholders and interested parties can view up-to-date progress toward IFAD's targets for the 79 indicators that Member States selected to track during the IFAD11 period. It is being adapted for IFAD12.

Independent evaluation

13. Independent evaluations are conducted by IOE, which is structurally, functionally and behaviourally independent of Management. IOE ensures that the entire evaluation function at IFAD follows internationally recognized good standards and practices. Independent evaluations help to reveal what has been achieved, what does or does not work and guide the development of successful policies and strategies to support rural transformation. The target audiences of independent evaluations are Management and governing bodies, Member States and the international development community at large.

14. IOE conducts a range of independent evaluations at different levels, including project, country programme, sectoral, thematic and corporate.

15. Project-level evaluations include independent validations of PCRs, project performance evaluations, impact evaluations and project cluster evaluations (the latter of which examine a set of projects in different countries that have a common topic of concentration – for example, rural finance). These products inform higher-plane evaluations, as well as the design of new and ongoing operations.

16. Country-level and regional evaluations include country strategy and programme evaluations (CSPEs) and subregional evaluations. CSPEs are usually conducted before IFAD and the government involved prepare a new results-based COSOP, and their findings and recommendations feed into the design of new COSOPs. Subregional evaluations assess intraregional issues or common development challenges within the region, in line with IFAD's decentralized business model.

17. Project- and country-level evaluations are the building blocks for evaluation syntheses. Evidence from past evaluations is synthesized and analysed to present evaluative knowledge on topics of strategic relevance and inform future directions and corporate-level evaluations. Corporate-level evaluations generate lessons and recommendations to improve IFAD's future policies and strategies. More detailed information and specific processes for each evaluation product are presented in part 2 of this manual.

18. Finally, the IOE's Annual Report on Results and Impact of IFAD Operations presents a synthesis of the performance of IFAD-supported operations and highlights systemic and cross-cutting issues, lessons and challenges that emerge from all independent evaluations. It also identifies recommendations to enhance IFAD's development effectiveness. As of 2022, IOE will produce the Annual Report on Independent Evaluation, a revamped version of its annual report.

II. Annex
Evaluations that are responsive to social justice and gender equality

1. **Social justice and no one left behind**. The commitment of the 2030 Agenda to "leaving no one behind"[1] needs to be reflected in evaluations. Evaluators should assess whether: (i) programmes have undertaken an analysis of the inequalities between different groups; (ii) the underlying drivers of such inequalities; (iii) whether programme designs address such inequalities; and (iv) whether the results frameworks of interventions have indicators to measure progress.[2] In line with a human rights-based approach, evaluations should assess the extent to which the initiative has facilitated the capacity of rights-holders to claim their rights and duty-bearers to fulfil their obligations.[3]

2. Typical examples of discrimination and power imbalance include: (i) economic factors (income, wealth); (ii) ethnicity (also indigenous status in some countries); (iii) social categories (including castes in some countries); (iii) gender; (iv) belonging to political groups/factions; (v) belonging to religious groups; and (vi) health and disability. This list is not comprehensive, and evaluators may identify other sources of discrimination and imbalance.

3. Evaluators need to assess how inclusive the intervention has been for different beneficiary groups and how key principles such as equity, non-discrimination and accountability have been incorporated from design to results.[4] There is a need to balance aggregation with specificity, with a sharper focus on the categories of the groups left behind and on "last mile" project delivery, rather than on average coverage and results. During data collection and data analysis, it is essential to consider the extent to which needs and priorities are being addressed. An analysis of differential results across groups and the extent to which the intervention contributes to or exacerbates equity gaps is a critical element for evaluations.[5]

1 The 2030 Agenda emphasizes the importance of empowering people who are vulnerable, including children, youth, persons with disabilities, people living with HIV/AIDS, older persons, indigenous peoples, refugees and internally displaced persons and migrants.

2 See United Nations Sustainable Development Group, "Leaving No One Behind: A UNSDG Operational Guide for UN Country Teams" (interim draft) (2019). https://unsdg.un.org/resources/leaving-no-one-behind-unsdg-operational-guide-un-country-teams

3 See United Nations Evaluation Group, Norms and Standards for Evaluation. New York: UNEG. Human rights and gender equality are considered a norm (Norm 8 on human rights and gender equality) and a standard (Standard 4.7, "The evaluation design should include considerations of the extent to which the United Nations system's commitment to the human rights-based approach and gender mainstreaming strategy was incorporated in the design of the evaluation subject.") (2016) http://www.unevaluation.org/document/detail/1914

4 United Nations Evaluation Group, "Guidance on Evaluating Institutional Gender Mainstreaming" (New York: UNEG, 2018), http://www.unevaluation.org/document/detail/2133

5 Evaluators should ensure the consistency and accuracy of terminology used in relation to gender issues in official documentation and publications, following the 2017 IFAD Glossary on gender issues: IFAD, Glossary on gender issues. (2017), https://www.ifad.org/en/web/knowledge/-/publication/glossary-on-gender-issu-1

4. **Gender equality as a specific area of attention for IFAD**. In line with the Fund's mandate, policies, strategies and work experience, evaluations at IFAD aim to be gender-responsive. This is accomplished by providing a systematic, impartial assessment that delivers credible and reliable evidence-based information about the extent to which an intervention has resulted in progress toward intended and/or unintended results in terms of gender equality and the empowerment of women. IFAD evaluations need to assess the degree to which gender and power relationships (including structural and other causes that give rise to inequities, discrimination and unfair power relations), change as a result of an intervention, using a process that is inclusive, participatory and respectful of all stakeholders (rights-holders and duty-bearers).[6]

5. At IFAD, performance in relation to gender equality and women's empowerment is assessed against a scale, moving progressively from "gender-blind" (i.e. there were no attempts to address gender concerns and/or the result had a negative outcome; aggravated or reinforced existing gender inequalities and norms) through "gender mainstreaming" (i.e. gender equality and women's empowerment have been mainstreamed, such that all three strategic objectives of the IFAD gender policy have been addressed),[7,8] all the way to, at the top of the scale, gender-transformative (i.e. gender power dynamics have been transformed by addressing social norms, practices, attitudes, beliefs and value systems that represent structural barriers to women's and girls' inclusion and empowerment). The meaning of "gender-transformative change" depends on the context (box 1). Different benchmarks are needed for different contexts, and good contextual analysis is a general prerequisite.

6 Evaluation approaches must integrate gender equality concerns and are all subject to assessment against the United Nations System-wide Action Plan on Gender Equality and the Empowerment of Women (UN-SWAP) **Evaluation Performance Indicator**. Likewise, all evaluations are subject to assessment against the **United Nations Disability Inclusion Strategy** (UNDIS), and the inclusion of persons with disabilities should be considered in all phases of the evaluation process and in every type of evaluation (**include hyperlink to IOE guidance note or UNEG guidance note - forthcoming*).

7 IFAD, "Policy on Gender Equality and Women's Empowerment," (2012), **https://www.ifad.org/documents/38711624/39417906/genderpolicy_e.pdf/dc871a59-05c4-47ac-9868-7c6cfc67f05c?t=1507215182000**

8 The three strategic objectives of the IFAD Policy on Gender Equality and Women's Empowerment are: (1) promote economic empowerment to enable rural women and men to have equal opportunity to participate in, and benefit from, profitable economic activities; (2) enable women and men to have equal voice and influence in rural institutions and organizations, and: (3) achieve a more equitable balance in workloads and in the sharing of economic and social benefits between women and men.

BOX 1

Definition of gender-transformative approaches[9]

Gender-transformative approaches are defined as those that aim to overcome the root causes of inequality and discrimination by promoting sustainable, inclusive and far-reaching social change. They actively seek to transform gender power dynamics by addressing social norms, practices, attitudes, beliefs and value systems that represent structural barriers to women's and girls' inclusion and empowerment. They seek to ensure equal access by women to productive assets and services, employment and market opportunities and supportive national policies and laws. Transformation and entry points toward it are context-specific and take into account that women are not a homogeneous group.

Source: Evaluation synthesis on gender equality (2017).

9 IFAD, "Evaluation Synthesis: What works for gender equality and women's empowerment - a review of practices and results), Independent Office of Evaluation (2017), **https://www.ifad.org/en/web/ioe/-/what-works-for-gender-equality-and-women-s-empowerment-a-review-of-practices-and-results**. **https://www.ifad.org/documents/38714182/39721405/gender_synthesis_fullreport.pdf/229358bf-f165-4dcd-9c4a-1af4f09ab065?t=1519897485000**; IFAD, "Report of the Consultation on the Twelfth Replenishment of IFAD's Resources: Recovery, Rebuilding, Resilience," (2020), **https://webapps.ifad.org/members/repl/12/4/docs/IFAD12-4-R-2-Rev-1.pdf?attach=1**

6. In assessing women's empowerment, evaluators may refer to the "domains of empowerment" outlined in the International Food Policy Research Institute (IFPRI)/FAO/IFAD Women's Empowerment in Agriculture Index guidelines, namely: decisions about agricultural production; access to and decision-making power about productive resources; control of the use of income; leadership in the community; and time allocation.[10,11] Evaluators may also refer to other analytical frameworks, such as the gender at work framework,[12] to better understand the types of changes that have taken place across the interlinked domains of individual change, formal change, systemic change and informal change.[13]

10 IFPRI, "Women's empowerment in agriculture index," (2012), https://www.ifpri.org/publication/womens-empowerment-agriculture-index
11 IFAD, "Measuring women's empowerment in agriculture: a streamlined approach," (2019), https://www.ifad.org/documents/38714170/39135332/19_Research_n%C2%B019_web.pdf/37a4a6ec-f23b-44da-8dea-8cefab20f295?eloutlink=imf2ifad
12 Analytical Framework – Gender at Work.
13 For more guidance see, for example, UN WOMEN, "Good Practices in Gender-Responsive Evaluations," (2020), https://www.unwomen.org/sites/default/files/Headquarters/Attachments/Sections/Library/Publications/2020/Good-practices-in-gender-responsive-evaluations-en.pdf

BOX 2

Example of issues to be explored that relate to gender equality[14]

- Volume and nature of project resources invested in gender equality and women's empowerment activities.
- Specific activities for gender equality and women's empowerment at the design stage.
- During implementation, the extent to which the project: (i) monitored gender-disaggregated outputs to meet gender-equality objectives; (ii) adapted implementation to better meet gender equality and women's empowerment objectives; (iii) addressed and reported on gender issues in supervision and implementation support; (iv) systematically analysed, documented and disseminated lessons on gender equality and women's empowerment; and (v) engaged in policy dialogue to improve gender equality and women's empowerment.
- In addition to others, changes in: (i) women's access to resources, land, assets and services; (ii) women's influence in decision-making; (iii) workload distribution among household members; (iv) women's health, skills, income and nutritional levels; and (v) gender relations within households, groups and communities in the project area.
- Changes in social norms, attitudes, behaviours, beliefs and value systems that represent structural barriers to women's and girls' inclusion and empowerment; and notice of whether such changes have been reflected in national policies and laws.

14 IFAD, "Evaluation Synthesis: What works for gender equality and women's empowerment - a review of practices and results," Independent Office of Evaluation (2017), https://www.ifad.org/en/web/ioe/-/what-works-for-gender-equality-and-women-s-empowerment-a-review-of-practices-and-results
https://www.ifad.org/documents/38714182/39721405/gender_synthesis_fullreport.pdf/229358bf-f165-4dcd-9c4a-1af4f09ab065?t=1519897485000
IFAD, "Report of the Consultation on the Twelfth Replenishment of IFAD's Resources: Recovery, Rebuilding, Resilience," (2020), https://webapps.ifad.org/members/repl/12/4/docs/IFAD12-4-R-2-Rev-1.pdf?attach=1

III. Annex
Core outcome indicators

TABLE A

List of core outcome indicators (to be collected by projects)[1]

	Core outcome indicator (COI)	Evaluation criteria for which COI is relevant
Access to natural resources	Core indicator (CI) 1.2.1: Households reporting improved access to land, forests, water or water bodies for production purposes	Impact Strategic objective (SO)1: productive capacities (agricultural/non-agricultural production and productivity) Improved nutrition
Access to agricultural technologies and production services	CI 1.2.3: Households reporting reduced water shortage vis-à-vis production needs CI 1.2.2: Households reporting adoption of new/improved inputs, technologies or practices CI 1.2.4: Households reporting an increase in production	Impact SO1: productive capacities (agricultural/non-agricultural production and productivity) Improved nutrition
Inclusive financial services	CI 1.2.5: Households reporting using rural financial services CI 1.2.6: Partner financial service providers with portfolio-at-risk ≥30 days below 5% CI 1.2.7: Partner financial services providers with operational self- sufficiency above 100%	Impact SO1: productive capacities (agricultural/non-agricultural production and productivity) Economic mobility
Nutrition	CI 1.2.8: Women reporting minimum dietary diversity (MDDW) (RMF 11) CI 1.2.9: Households with improved knowledge, attitudes and practices (KAP)	Impact SO1: productive capacities (agricultural/non-agricultural production and productivity) Improved nutrition
Diversified rural enterprises and employment opportunities	CI 2.2.1: New jobs created (IFAD11) In IFAD12, this indicator will be substituted with IFAD12 RMF indicator: Beneficiaries with new jobs/employment opportunities CI 2.2.2: Supported rural enterprises reporting an increase in profit	Impact SO2: access to market (access and integration into markets) Impact Economic mobility
Rural producers' organizations	CI 2.2.3: Rural producers' organizations engaged in formal partnerships/agreements or contracts with public or private entities CI 2.2.4: Supported rural producers' organizations reporting new or improved services provided by their organization CI 2.2.5: Rural producers' organizations reporting an increase in sales	Impact SO2: access to market (access and integration into markets)
Rural infrastructure	CI 2.2.6: Households reporting improved physical access to markets, processing and storage facilities	Impact SO2: access to market (access and integration into markets)

1 CI Guidelines, October 2021.

	Core outcome indicator (COI)	Evaluation criteria for which COI is relevant
Environmental sustainability and climate change	CI 3.2.1: Greenhouse gas emissions (CO2e) avoided and/or sequestered (RMF11) In IFAD12, this indicator will be substituted with IFAD12 RMF indicator: Tons of greenhouse gas emissions (tCO2e) avoided and/or sequestered CI 3.2.2: Households reporting adoption of environmentally sustainable and climate-resilient technologies and practices (RMF11) CI 3.2.3: Households reporting a significant reduction in the time spent for collecting water or fuel	Impact SO3: resilience (ability of households to cope with climate and non-climate shocks)
Policy	Policy 3: Existing/new laws, regulations, policies or strategies proposed to policymakers for approval, ratification or amendment	Sustainability Policy engagement
Empowerment	Individual empowerment 2.1: Individuals demonstrating an improvement in empowerment	Sustainability
Stakeholder feedback (SF)	SF 2.1: Households satisfied with project-supported services SF 2.2: Households reporting they can influence decision-making of local authorities and project-supported service providers	Sustainability

TABLE B

Mandatory CIs, by project type[2]

Mainstreaming themes and corporate commitments	Related indicators	Use and requirements
Climate finance - Adaptation	**OUTCOME** CI 3.2.2: (Number) Percentage of persons/households reporting adoption of environmentally sustainable and climate-resilient technologies and practices CI 3.2.3: (Number) Percentage of persons/households reporting a significant reduction in the time spent for collecting water or fuel	At least one of the following CIs. The higher the share of adaptation finance, the more intervention-appropriate indicators may be selected.
Climate finance - Mitigation	**OUTCOME** CI 3.2.1: Tons of greenhouse gas emissions (tCO2e) avoided and/or sequestered.	If appropriate, mandatory
Gender-transformative	**OUTREACH**: Disaggregated by sex **OUTCOME**: CI IE.2.1: Individuals demonstrating an improvement in empowerment	Mandatory Mandatory
Nutrition-sensitive	**OUTREACH**: Disaggregated by sex and youth **OUTCOME**: CI 1.2.8: Percentage of women reporting minimum dietary diversity (MDDW14) CI 1.2.9: Percentage of households with improved nutrition KAP	Mandatory Mandatory At least one outcome CI mandatory
Youth-sensitive	**OUTREACH**: Disaggregated by sex and youth	Mandatory
Stakeholder feedback	**OUTCOME**: CI SF.2.1: (Number) Percentage of households satisfied with project-supported services CI SF.2.2: (Number) Percentage of households reporting they can influence decision-making of local authorities and project-supported service providers	Both mandatory in project logframes approved from December 2020 onwards

2 CI Guidelines, October 2021.

IFAD

Investing in rural people

Independent Office of Evaluation
International Fund for Agricultural Development
Via Paolo di Dono, 44 – 00142 Rome, Italy
Tel: +39 06 54591 – Fax: +39 06 5043463
E-mail: evaluation@ifad.org
www.ifad.org/evaluation
www.twitter.com/IFADeval
www.youtube.com/IFADevaluation

ISBN 978-92-9266-265-3

9 789292 662653