Food and Agriculture Organization of the United Nations

USING ARTIFICIAL INTELLIGENCE TO ASSESS FAO'S KNOWLEDGE BASE ON THE TECHNOLOGY ACCELERATOR

USING ARTIFICIAL INTELLIGENCE TO ASSESS FAO'S KNOWLEDGE BASE ON THE TECHNOLOGY ACCELERATOR

Published by
Food and Agriculture Organization of the United Nations
Rome, 2023

Required citation:

FAO. 2023. *Using artificial intelligence to assess FAO's knowledge base on the technology accelerator*. Rome.
https://doi.org/10.4060/cc6724en

ISBN 978-92-5-137967-7

Graphic designer: Laura Monopoli

Cover photograph: © FAO/Mohammad Rakibul Hasan, © FAO/FameMedia, © FAO/Gideon Vink

Contents

FIGURES

TABLES

BOXES

Abbreviations and acronyms

AI	artificial intelligence
ATIO	Agrifood System Technologies and Innovations Outlook
BERT	bidirectional encoder representations from transformers
CRISPR	clustered regularly interspaced short palindromic repeats
DNA	deoxyribonucleic acid
FSN	Global Forum on Food Security and Nutrition
GM	genetic modification
GPT	generative pre-trained transformer
IFC	International Finance Corporation
LLM	large-language model
IoT	Internet of Things
ML	machine learning
MLMs	machine-learning models
NAL	United States National Agricultural Library
NER	named entity recognition
NLP	natural language processing
REDD	Reducing Emissions from Deforestation and Forest Degradation
SMEs	small and medium-sized enterprises
SDGs	Sustainable Development Goals
STI	science, technology and innovation
UNFCCC	United Nations Framework Convention on Climate Change
URL	uniform resource locator
WEAI	Women's Empowerment in Agriculture Index
WHO	World Health Organization

Acknowledgements

The report was prepared by a multidisciplinary team led by FAO and the University of Notre Dame, under the direction and coordination of Preetmoninder Lidder, Technical Adviser, Office of the Chief Scientist, and overall supervision of Mona Chaya, Special Adviser, Office of the Chief Scientist.

University of Notre Dame: Jaron Porciello, Shamaila Ashraf, Volha Skidan, Paul Winters and Dakota Peterson.

Additional inputs were provided by the following FAO colleagues:

Norberto Ambagan, Anne Aubert, Sergio Bogazzi, Henry Burgsteden, Olivier Dubois, Carola Fabi, Jippe Hoogeven, Josef Kienzle, Christian Mongeau, Maryam Rahmanian, Imma Subirats and Jonathan Wirths.

Special thanks go to Jonathan Robinson (FAO) for editing of the text, Ludovica Mei (FAO) for production coordination and Laura Monopoli for the design and final layout of the publication.

Executive summary

Supporting the goals of efficient, inclusive, resilient and sustainable agrifood systems requires, among other levers of change, that an ensemble of different technologies work together to achieve the needs of a diverse and ever-changing world. Farm productivity must be sufficient to feed a growing global population, even as climate-change-induced droughts, heavy rains and high temperature events are increasing in number, intensity and unpredictability. Digital technologies, including remote sensing, can help safeguard water use and minimize waste in areas of water scarcity. Mechanization and automation can offer substantial advantages in terms of increased precision, reduced drudgery and more efficient allocation of labour. Other technologies, including biotechnologies, can support a global shift towards healthier diets as well as expanding the array of options that small-scale producers have to adapt to and mitigate climate change. There are also important innovations throughout the entire value chain, ranging from food processing technologies to renewable energy technologies, which respond to an increased demand to meet national and global climate targets.

The Food and Agriculture Organization of the United Nations (FAO) 2022–2031 Strategic Framework highlights a series of cross-cutting opportunities to maximize progress in all of its programmatic interventions. The four "accelerators" – technology, innovation, data, and complements (governance, human capital and institutions) – are paramount to fostering rapid progress across agrifood systems.

Technology holds significant opportunities to transform agrifood systems and improve food production, while minimizing negative impacts on the environment. The FAO State of Food and Agriculture 2022 report emphasizes the role of automation in agricultural development, highlighting that technological innovation has a long history of delivering gains across the sector. At the same time, the report also notes that automation can be a source of unintended consequences that exclude some communities and empower others, describing the barriers that technology can create. In order to ensure that agricultural automation is inclusive, a multifaceted approach is necessary, encompassing policies, investments, and interventions on various fronts.

This report examines the technology accelerator trends across publicly available FAO knowledge reports, technical guidance and convening summaries. Leveraging AI-assisted classification of nearly 40 000 documents, this report offers a bird's-eye perspective of six types of technology - digital technologies, biotechnologies, mechanization, irrigation technologies, renewable energy technologies and food processing technologies - as well as high-level trends for outcomes and social and demographic details about the communities using these technologies.

These technology areas are intentionally broad. They represent cross-sectoral areas that can be beneficial for increased research, development and investment to both public and private sector actors across the agrifood system, including small-scale producers and small-and-medium sized enterprises, as they seek to address multiple, and sometimes competing, objectives of sustainable production of nutritious food.

The findings highlight some noteworthy trends. Knowledge and information about digital technologies and biotechnologies is well represented, while information on renewable energy and food processing technologies is less so. Given the challenges posed both by climate change and disruptions to global supply chains, and increasing poverty and inequalities exacerbated by the COVID-19 pandemic and ongoing conflicts, the dearth of knowledge and solutions from underrepresented areas could have significant effects. For example, increased knowledge and awareness about scaling food processing technologies at the subnational level can promote opportunities to reduce post-harvest food loss and waste and increase off-farm economic growth year-round for communities. It may also provide insight into increased opportunities to improve dietary diversity if such technologies emphasize processing a diverse set of underutilized and nutritious crops that reflect local preferences. Similarly, increased knowledge and specific, actionable examples about solar power, wind, geothermal, biomass and other renewable energy sources are critical for farmers and other market actors to accelerate their transition away from fossil-fuel-based technologies.

The analysis incorporates a trend analysis about outcomes. Outcomes are the cornerstone of evidence-informed assessments and decision-making. Outcomes provide context about the high-level aims that any project or programme seeks to achieve, and they are important for the development community to seek out and reflect upon in order to ensure that progress towards global goals is reflected in real and actionable ways by the larger agrifood systems community. Outcomes related to economic growth, such as incomes and productivity, and food security and nutrition, were most mentioned across all publications, but areas that are high priority, including environmental outcomes and social inclusion, were some of the least mentioned.

This report also explores social and demographic details about the populations of people who are impacted both directly and indirectly by technologies. The state of evidence is assessed in FAO publications supporting how populations are described and represented. Data on small-scale producer populations are both under-reported and inadequately described. Key details about sex, age, wealth and education – essential data to understand how to maximize and improve technology use and efficiency – are largely absent.

Significant investments are being made in knowledge production, information technology, data curation and digital platforms across FAO. Investigating new ways of optimizing the data means that the value of the data can be extended further to deliver greater insights. Advanced analytics, supported by innovative approaches, can be important for improving stakeholder engagement interaction, and draw on evidence-based assessments in complex, data-rich environments.

Equitable deployment of technologies requires understanding the barriers and facilitators for accessing technologies for different populations and communities. However, there is no standard monitoring, tracking and assessment of technology use and impact among small-scale producers, despite increased attention paid to inclusivity and resilience. FAO has been studying technology use for decades and is in a unique position to produce thoughtful leadership on this issue, including through a new knowledge product - the Agrifood System Technologies and Innovations Outlook (ATIO). As technologies continue to evolve and transform, FAO will work with partners to identify promising technologies and support countries to access the latest and most appropriate solutions, adapted to their contexts.

This report's landscape analysis across publications and information that FAO has produced helps identify strengths and gaps in FAO's dissemination of technology for agrifood systems as well as determine trends in technologies that FAO should monitor in the coming years. By creating systematic approaches to analyse the breadth and depth of knowledge and guidance across key technology and document indicators, FAO has the potential to use this assessment to help set regional and national priorities, inform future interventions, and identify gaps in knowledge and guidance provided.

Introduction

Harnessing science, technology and innovation (STI) is key to meeting the aspirations of efficient, inclusive, resilient and sustainable agrifood systems and leveraging emerging opportunities to achieve the Sustainable Development Goals (SDGs). As the lead United Nations specialized agency for food and agriculture, FAO is at the forefront of facilitating solutions that support the transformation of agrifood systems for better production, better nutrition, a better environment and a better life, leaving no one behind.

Technology is an instrumental part of the package of solutions needed to transform agrifood systems, and the development and diffusion of technologies and associated knowledge can be a powerful driver of sustainable development. The 2019 Global Sustainable Development Report identified "science and technology" as one of the levers for transformation to accelerate progress in achieving the SDGs and minimize trade-offs (United Nations, 2019). At the same time, technologies are not neutral. They are embedded in, and have influence on, social and economic relations. This leads to potential trade-offs and unintended consequences, especially in the long term.

The FAO Strategic Framework 2022–2031 identifies STI as having enormous transformative potential and underlines the potential of emerging technologies (sometimes referred to as the fourth industrial revolution). It also recognizes that STI can present substantial risks, such as reinforcing inequality and market concentration, or contributing to the degradation of natural resources (FAO, 2021a).

The FAO Science and Innovation Strategy, a key tool to support the delivery of the FAO Strategic Framework 2022–2031, aims to ensure that FAO Members harness science and innovation to realize context-specific, sustainable and systemic solutions for MORE efficient, inclusive, resilient and sustainable agrifood systems across the four "betters"—better production, a better environment and a better life leaving no one behind, in support of the 2030 Agenda for Sustainable Development. The Strategy includes several guiding principles that can help to ensure that the development and use of science, technology and innovation contribute to the values of the 2030 Agenda for Sustainable Development (FAO, 2022a).[1]

1 The guiding principles are rights-based and people-centered, gender-equal, evidence-based, needs-driven, sustainability-aligned, risk-informed and ethics-based.

As one of four accelerators identified by the FAO Strategic Framework 2022-2031,[2] technology is expected to "accelerate impact while minimizing trade-offs". The FAO Strategic Framework 2022-2031 takes a systemic approach that seeks to overcome siloed thinking by considering the four accelerators in relation to each other, while also ensuring that they address the cross-cutting themes of gender, youth and inclusion.

FAO plays an important role in supporting the introduction of technologies in countries, as well as the sharing of knowledge and technologies among countries. In doing so, FAO must take care to assess technologies in the broad context of agrifood systems challenges in all their dimensions, recognizing that under specific conditions technologies can address challenges of hunger and malnutrition, and in others they can be either neutral or may even enhance problems. Having a harmonized understanding of how and where technologies have evolved, including the prominence that certain technologies have gained, allows FAO to evaluate its strategic priorities, in addition to examining the role that technology has played in agrifood systems transformation.

To understand better the existing organizational knowledge and resources on technologies, this report analyses the breadth and depth of work on the topics published by FAO, allowing for a baseline assessment of FAO's role to provide guidance and knowledge on development of key technologies and their use in the agrifood sector. At the same time, this analysis pays equal attention to other dimensions, such as the relationship of technology to development outcomes and the people and populations that are referenced.

The use of artificial intelligence and machine-learning models (MLMs) can save organizations and end-users of research and data valuable time in summarizing and benchmarking data that can be put to greater use. New technologies around large-language models (LLMs) and natural language processing (NLP) offer promising opportunities to explore rich scientific and technical text-based content with the help of algorithms. Such efficient machine-learning approaches, when applied intelligently and using relevant data, can be a highly effective way of revealing relevant insights from a large and representative dataset (Gil *et al*., 2014).

Organizations such as FAO produce significant volumes of knowledge and data each year. Gartner Research reports from 2021 estimate that nearly 80 percent of the world's information exists as unstructured data (Panetta, 2021). Working with unstructured data, such as textual data, requires more sophisticated approaches for data cleaning, analysis and summarization than working with structured big data, such as weather or satellite data (McCallum, 2005). Increasingly, more organizations are paying attention to the value of their own unstructured data and are leveraging AI-based approaches to foster deeper, more comprehensive understanding on particular issues (Faccia *et al*., 2022). While academic and development sectors still lag behind the private sector in using AI and ML approaches to explore organizations' data to improve decision-making, more organizations, including development and funder organizations in agriculture, are realizing the value that textual analysis can offer to sequence and isolate millions of data points over time from project summaries, reports and other datasets to assess how they interact with each other.

2 The four accelerators include technology, innovation, data, and complements (governance, human capital, and institutions).

The FAO Strategic Framework 2022-2031 notes both the use of data and advanced technologies to optimize their use, including modernization of policies and business models for agrifood systems. Better use of AI and MLM enables organizations to review large amounts of data stored across multiple teams and repositories (Garbaro *et al*., 2020; Porciello and Ivanina, 2021). Such approaches help to reveal meaningful patterns and insights across data that are otherwise unable to be synthesized, except through manual review.

Artificial intelligence, including LLMs and NLP, continues to advance at a rapid pace. However, ML models must be calibrated during training for specific tasks and training data must be accurate and relevant for the tasks. Data diversity, high-quality training data, acknowledgement of data sources and of underlying models are all important factors that can help reduce the risks associated with using AI to support knowledge exercises.

This exercise contributes a landscape analysis that uses AI across publications and information that FAO has produced, identifies different uses of technology, and tracks the density and frequency of knowledge published by FAO. It also helps identify strengths and gaps in FAO's dissemination of information on technologies for agrifood systems as well as determine trends in technologies that FAO should monitor in the coming years. By creating systematic approaches to analyse the breadth and depth of knowledge and guidance across key technology and document indicators, FAO has the potential to use this assessment to help set regional and national priorities, inform future interventions, and identify gaps in knowledge and guidance provided.

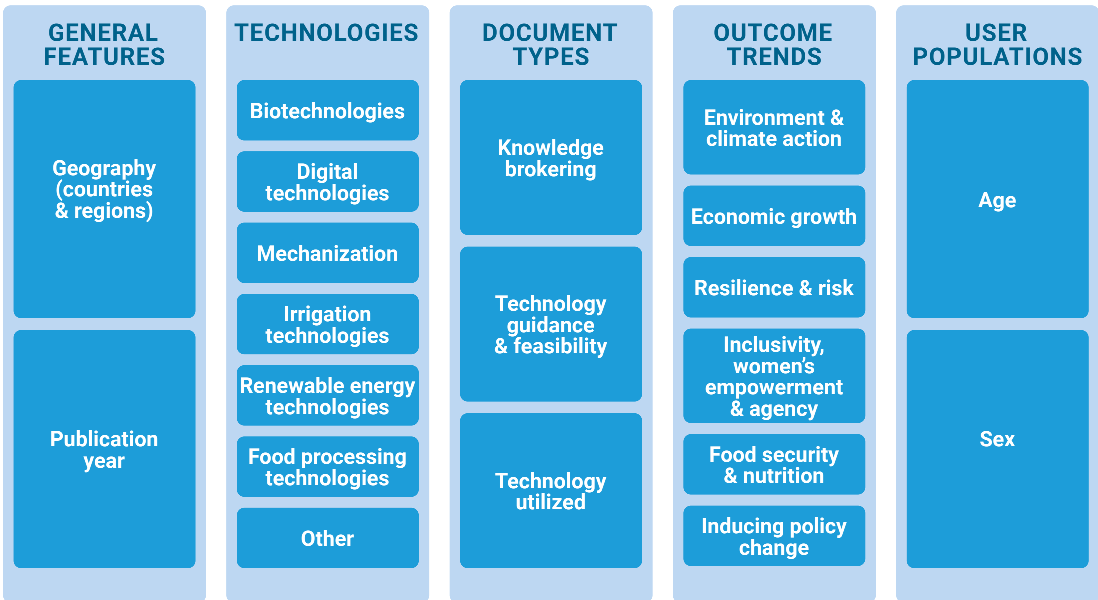
Defining the scope

This project was designed to take stock of the publicly available publications and guidance produced by FAO on key areas of technologies. This type of assessment is designed to identify gaps by evaluating the breadth, and to a more limited extent, the depth, of what can be learned from the available evidence using a series of standardized indicators.

Using over 39 000 relevant documents from the FAO document repository, the contents were summarized according to a typology (Figure 1) that, for the purpose of this exercise, included six key technology types and one other, the type of knowledge, guidance, or information they provided, and any available summary information about user groups, geographies and outcomes.

FIGURE 1

Project typology, including the technology types that were evaluated



Creating the typology and conducting the analysis was split into three phases. Phase one included a series of workshops and exercises to define priorities, key terms, and explore the breadth of expertise that should be integrated into the data curation process. Phase two focused on identifying the most comprehensive, complete and correct set(s) of data available within FAO repositories that could be used to inform the typology. Phase three focused on the ML and textual analysis of the data to extract key information from the publications to fill the typology and provide analysis on the overall trends of the technology accelerator.

Technology types

The six technology areas are intentionally broad and can be further refined over time. Currently, they represent a series of cross-sectoral areas that can be beneficial for increased research, development and investment to both public and private sector actors across agrifood systems, including small-scale producers and small and medium-sized enterprises (SMEs), as they seek to address multiple, and sometimes competing, objectives of sustainable production of nutritious food. Identifying the six technology areas for this pilot study was a process that involved expert knowledge working groups and exploration of the available assets on the technology accelerator. This included resources that were developed as part of an internal mapping exercise, drawing on important baseline data, including expert-defined definitions (Glossary) and keywords (Annex A).

Knowledge categories

Given that the analysis looked at more than 39 000 documents published between January 2008 and June 2022, a clear way of categorizing documents was needed so that information could be processed quickly and to understand the way in which technology has been addressed across documents. The documents of interest included reports, convening results, conference findings, e-consultation documents, toolkits, project evaluations, books and policy briefs. Subscription content or content from other organizations (such as World Bank documents) was not included. Only documents published by FAO or translated by FAO within the repository in English were considered for this study. Processing materials in other languages was outside the scope of this study, though it is recommended for future iterations of this exercise.

The dataset from the FAO Publications Workflow System included the following document types: book (stand-alone), book (series), brochure, flyer, factsheet, booklet, policy brief, journal, magazine, bulletin, newsletter, infographic, meeting, project, document and article. Casting a wider net of the FAO Publications Workflow System document labels allowed the inclusion of non-traditional document sources, such as bulletins, for this pilot phase. All citations included a URL (uniform resource locator) to access the full text.

Document categories (Figure 2) were created through an iterative process, including validation and adjustment process across the technology types by a small panel of experts. This classification represents a generalized framework that can be updated and improved over time, and importantly, was created *ex post,* i.e. after publication of the materials. Draft categories were presented with definitions of the types of documents fitting into a category based on a manual search of different document types in FAO repositories.

The three primary document type categories used for this study were:

- **Knowledge brokering**. FAO plays an important role in knowledge brokering, where information about current, new and emerging technologies, their impacts and use cases, are collected from the community through convenings, background knowledge and expert assessments, evidence reviews and other research or technical activities. The culmination of these activities allows FAO to act as a hub for information and knowledge to support the broader community. Knowledge brokering documents typically include flagship reports, such as the State of Food Security and Nutrition, as well as expert commissioned pieces such as a recent publication on gene editing and agrifood systems (FAO, 2022b).

- **Technical guidance and feasibility**. Acting as a collaborator with local and national entities, FAO offers coordinated input based on information from standing committees and country-level dialogues to address various strategic priorities and initiatives. This includes applied research methods, local entity collaborative efforts that FAO facilitates or contributes to. Technical guidance and feasibility documents include summaries produced as a result of expert consultations, including the Global Forum on Food Security and Nutrition (FSN Forum).

- **Technology utilized**. Technologies that have been utilized or tested as part of FAO projects across the existing categories. Examples and use cases about these technologies can be found within knowledge brokering and technical guidance and feasibility materials. Examples can also be found within other data sources, including financial costs and impact and field innovation reports. Technology-utilized documents include project reports that often provide real-world context and insight into the facilitators and barriers to technology implementation.

FIGURE 2

Document category types based on expert elicitation exercises

KNOWLEDGE BROKERING	TECHNICAL GUIDANCE	TECHNOLOGY UTILIZED
Relevant documents: notes and agendas from forums, events and convenings, dialogues, expert analysis documents	**Relevant documents:** formal consultations, country requests, policy guidance reports	**Relevant documents:** project reports, sources and documents where technology implementation has been highlighted
Subcategories: • Convening • Analysis	**Subcategories:** • Tech resources & guidelines • Feasibility • Assessment materials • Capacity strengthening	**Subcategories:** • Use of third party technology • Field innovations • Financial costs and impacts
Example keywords: meeting notes, convening, conference, convene, report, meeting, working group, session, analysis, review, overview, brief, ecosystem, proceedings, agenda, minutes	**Example keywords:** guidance, resource, guide, guidelines, stakeholder mapping, needs assessment, assessment, database, strategic plan, plan, strategy, manual, appraisal	**Example keywords:** technical report, dataset, data, experience, field, field study, financial, evaluation, results, result, pilot

Data identification and curation

Once clear definitions had been created and validated, phase 2 involved mapping out FAO's data landscape and gaining access to a variety of datasets to complete the analysis. Box 1 provides an overview of the sources in the dataset and how it was assembled.

Initial review of the dataset offered the following insights as to how FAO publications are organized. Of the 39 579 documents, only 10 436 contained abstracts or summary information (Figure 3).

Document titles rarely provided enough information for content analysis. This limiting feature, and other data curation challenges encountered during the process, are described in Box 1.

Most of the documents were either in English, or had been translated into English (title, subtitle, summary where applicable) through the FAO Publications Workflow System (Figure 3). The parameters for analysis were documents published from 2008 onwards that demonstrated steady growth across nearly all years (Figure 4).

FIGURE 3

The FAO Publications Workflow System export provided metadata about FAO publications for title, subtitle and abstract. The data show an imbalance of available metadata at the document level

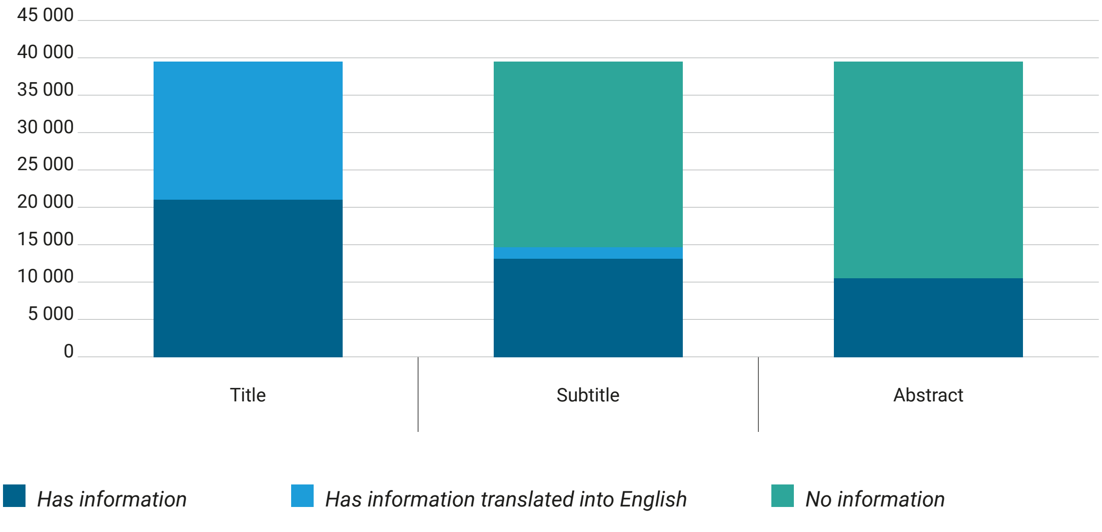
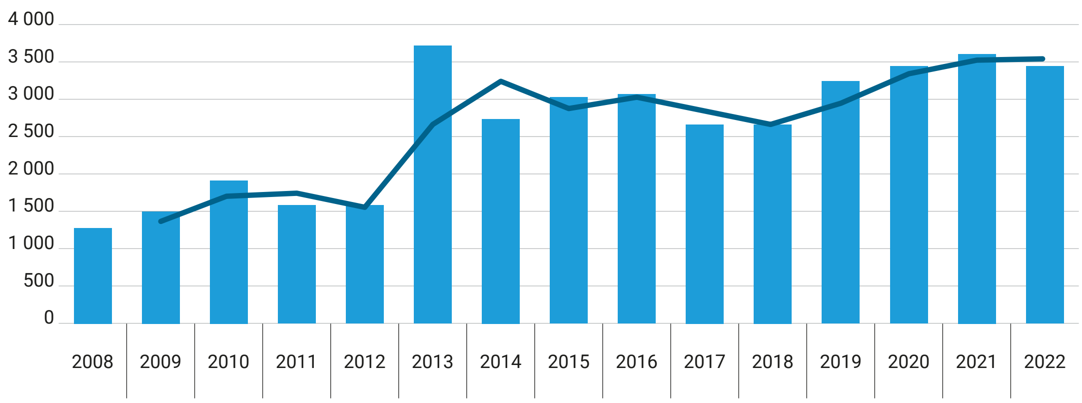


FIGURE 4

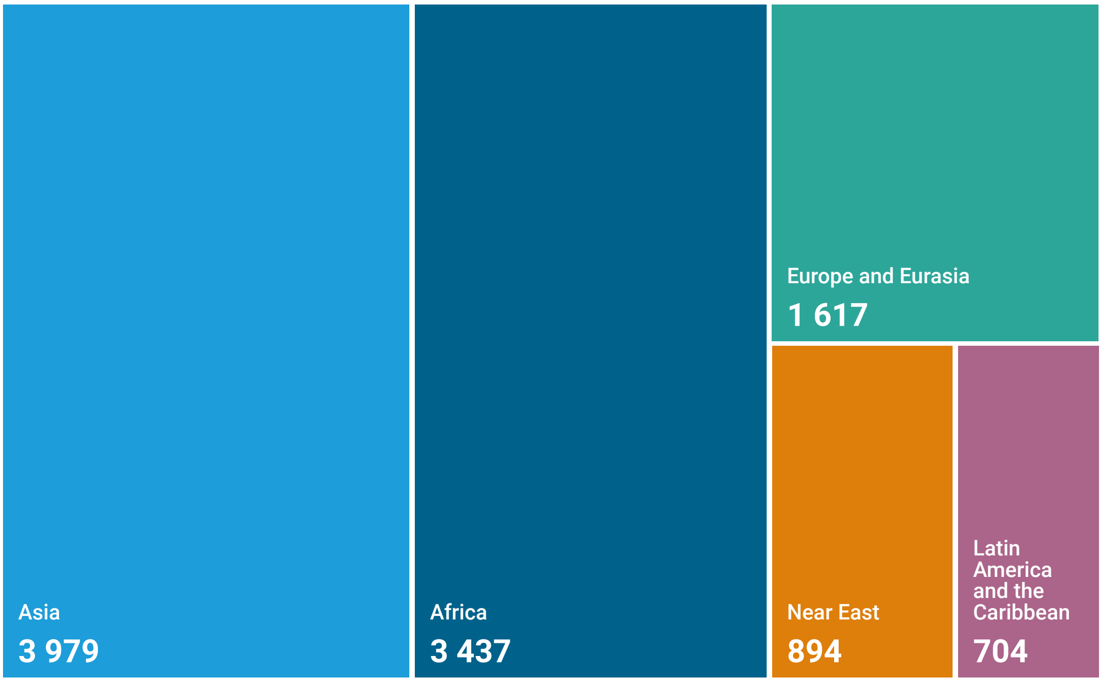
The FAO Publications Workflow System publication count by year



Because one of the parameters for data extraction was that documents should be published by FAO directly, very little information can be obtained by looking at publication locations. However, a process to extract and label country and region from document summaries shows how publications are spread based on geography (Figure 5), which indicates that regional mentions were highest for Asia and Africa, while Latin America had one of the lowest mention rates in this dataset. This is possibly due to English language requirements of the study. However, other landscape studies looking at thousands of documents on both livestock interventions and policy interventions, and where Spanish language materials were incorporated, reported similar gaps (Baltenweck *et al.*, 2020; Piñeiro *et al.*, 2020).

FIGURE 5

FAO regional mentions



BOX 1

Assembling the dataset for curation: challenges and opportunities

Data curation is the process of creating, organizing and maintaining datasets for better access and analysis of information. A challenge with data curation is interoperability – or the ability to combine data from one system with data from another. Given that no single system provides access to all the world's data, access and interoperability are important considerations for a large-scale data analysis project.

One of the elements used to support data curation is metadata. Metadata are data that describe information about other data, making it easier to find relevant data within large repositories. For instance, when there is a required metadata field to include "geography", high-quality metadata would use a controlled vocabulary of countries, geographical regions etc., so that the data are always standardized. Low-quality metadata, on the other hand, might still provide a field for a country name but without standardized vocabulary, making them vulnerable to errors.

Working directly with the FAO knowledge management teams to obtain direct access to datasets was an important component for the success of this analysis. As in most large organizations, no one single system at FAO provided all the data needed for this analysis. In addition, each system presented some unique technical challenges that required additional technical work. Such challenges are common in a large, knowledge-rich organization, where systems and processes have developed over time to meet the changing needs of the organization.

For instance, the FAO Document Repository, which provides discovery of FAO publications, currently lacks some user interface features to select and download multiple citations resulting from a search. Without this feature, it is challenging to obtain a list of underlying citations and their available metadata from a search query. The FAO Document Repository is being migrated to a new platform that will offer advanced functionalities for harvesting and discovering metadata and full text documents, which will help facilitate enhanced knowledge and discovery about agrifood systems. The new platform will be ready in 2023.

The FAO Agricultural Science and Technology Information database, which currently indexes metadata for more than 14 million books, journal articles, monographs, book chapters, datasets and grey literature from the 1970s to the present – including unpublished scientific and technical reports, theses, dissertations and conference papers in the area of food and agriculture – does not yet provide a comprehensive set of all FAO documents, but work is ongoing to improve the availability of FAO documents. In a sample dataset of more than 200 000 records, fewer than 10 percent of the citations from 2007 onwards included a URL to the corresponding online document. This is an important element to conduct full-text document analysis. On the other hand, FAO Agricultural Science and Technology Information database data are indexed using the FAO multilingual thesaurus (FAO, 2023). The FAO multilingual thesaurus is a Linked Open Dataset about agriculture available for public use and facilitates access and visibility of data across domains and languages. It offers a structured collection of agricultural concepts, terms, definitions and relationships that are used to identify resources unambiguously, allowing standardized indexing processes and making searches more efficient. It uses semantic web technologies, linking to other multilingual knowledge organization systems and building bridges between datasets.

The FAO Publications Workflow System has been designed to help FAO divisions and offices plan and manage their knowledge products, from conceptualization to publishing, with checkpoints for internal clearance and quality control from the Publications Branch, for final approval. It is a monitoring tool that contributes to the production of cost-effective, high-quality and targeted publications. At the very beginning of the preparation process, the FAO Publications Workflow System prompts originators to reflect on the purpose and costs of their publishing products, and their alignment with the FAO Strategic Framework 2022-2031 and SDG indicators.

The FAO Publications Workflow System includes different document types, from the most prestigious publications, such as flagships, to grey literature, meeting documents and presentations. It provides an overview of all publishing activities within the Organization to facilitate analysis and strategic planning of FAO's publishing programme.

However, considering the large diversity of document types, not all the documents receive the same treatment in terms of metadata. There are comparatively limited summary data in the FAO Publications Workflow System because of the volume of data that FAO manages each year (detailed data are provided only for a subset of publications). Only about 25 percent of documents had an abstract (Figure 3). The FAO Publications Workflow System metadata also had inconsistencies between presence of English language, or English translations in the title, that required data cleaning. The curation process of the metadata focusses on specific product types (such as books, booklets and policy briefs) and does not include many details for product types such as posters and infographics.

Given these types of challenge, introducing data science tools and approaches for rich textual analysis is increasingly important. It provides opportunities to extract better insights from existing publications and reports. One approach that was taken was to look at the crossover between citations that had FAO Agricultural Science and Technology Information database metadata, and the FAO Publications Workflow System citations with no abstract, to establish if there were overlaps and insights that could be extracted. This proved successful and provided some relevant training data for the phase three model analysis.

Source: **FAO**. 2023. AGROVOC. In: *Food and Agriculture Organization (FAO) of the United Nations*. Rome. Cited 21 February 2023. https://www.fao.org/agrovoc/

Using artificial intelligence to support document analysis

The use of deep learning LLMs that feature advanced NLP functionality are useful because they can perform tasks related to textual analysis with speed and accuracy. One of the most recent LLMs, generative pretrained transformer 4 (GPT-4) performs well for generalized questions, answers and text completion. The bidirectional encoder representations from transformers (BERT) used in this study are extremely effective at learning contextual relationships between words and sentences (Devlin *et al*., 2018).

ML offers a useful approach to support identification of patterns and becomes more dynamic over time as the models are exposed to more data and information. As with any technology, there are limitations and some risks associated with using AI. The volume and type of training data significantly factors into how well the model will perform. Too little variability in the training data will create a situation known as overfitting the model, where an inaccurate label is produced when new data are introduced into the model because it has been exposed to too little data. More information about ML training and assessment is provided in Box 2.

This pilot introduces a hybrid human machine approach for classification and review of documents but does not provide autonomous recommendations for decision-making. An ML pipeline supported the detection and labelling of technology types (single label), document types (multi-label), and evaluation of outcomes mentioned (multi-label). The training data were assembled from scientific and technical documents, including the FAO summary data, and were collaboratively coded by the research team and FAO experts.

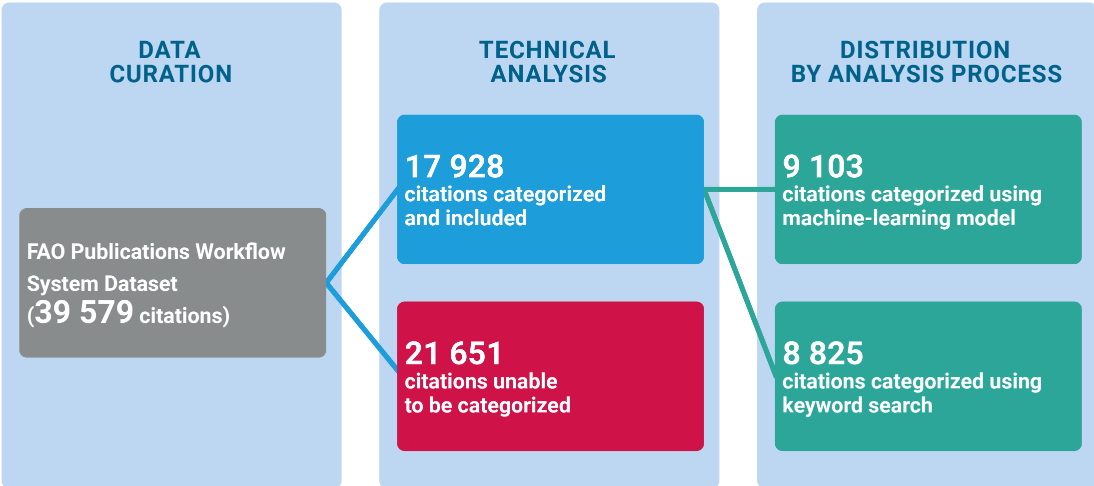
Given that the title data exceeded abstract data (Figure 3), a two-pronged approach was used to support labelling documents with a primary technology label and multi-label document category label (co-occurrence of document category labels is available in Annex B). A numerical value between zero and one was produced for technology type, known as a confidence score, to provide thresholds for the model outputs. A process to identify additional keywords associated with each technology category was designed in order to create a prioritization process to accept a keyword-based label or a machine-based label for technology types when the document had only title data available. If a machine-modelled threshold of 0.75 was not achieved for documents with title data only, the document was categorized using a keyword function. If no keyword was identified from the expanded synonym list, a label of “no technology found” was applied. The threshold of 0.75 was identified as an accurate threshold during training data and expert review phases. Additional details are provided in Annex B including Table 3, which provides information about the performance of the models.

One caveat to this process that is discussed in the Technology Acceleration section is mechanization labels. Insufficient data were associated with this technology type. For this category, if a label was applied with the keyword search and the model returned a confidence score of 20 percent or higher, FAO used the ML model.

About 45 percent of articles were classified using this hybrid approach – totalling 17 928 documents categorized. The remaining 21 651 articles were not labelled because a technology (based on the categories determined by FAO) was not found in the document data (title, subtitle, or abstract). This process is presented in Figure 6.

FIGURE 6

Technical process to analyse and classify the dataset.
The challenges of summary and metadata availability that contributed to a two-pronged process for analysis are described in Box 1



In addition to looking at technology types and document information, further evaluation of a subset of documents that presented an abstract (10 436, Figure 3) was conducted to identify outcome trends and explore social and demographic details about individuals and communities. These results are indicative of trends only. A more robust analysis that included evaluation of the full text is needed to capture important information that was not presented in the abstract. Such assessment is much more time and resource intensive.

BOX 2

Machine learning: an overview

Machine learning tasks require training to improve overall accuracy. The training is typically conducted in the form of humans providing small amounts of feedback by labelling data. For models to enhance their ability to operate independently, they must be trained.

Large-language models like BERT and GPT-4 are pretrained with general language and concept understanding through exposure to Google News, Google Images and Wikipedia, so that they can readily execute many tasks without additional training.

However, pretrained models must then be fine-tuned to perform other tasks, such as language inference about scientific and technical concepts within specific datasets. This is typically done using data that humans have reviewed and labelled data for certain characteristics. In the simplest example, human reviewers can label photos as "apple" or "orange" before a data scientist runs a series of tests to determine accuracy of the model to correctly pick the right fruit for the right label. Once the model performs the task accurately, it can be done without human input. This is the process of machine learning.

Training for the models leveraged for this analysis has occurred across different use cases in agriculture (Porciello *et al*., 2020; 2022). A primary feature of the models is the ability to identify, classify and cluster agricultural interventions and outcomes into a relevant taxonomy.

This model is unique in its ability to capture concepts instead of relying on keywords associated with certain categories.

How is machine learning assessed?

To test the accuracy of the process, data are extracted from the model and randomly split into batches; some of the data are reviewed and corrected, others are held aside for testing.

Measuring model performance (accuracy) is conducted using both precision and recall. Precision is the model's ability to express the proportion of the data points that it says were relevant and it is measured by assessing how many of the selected items are relevant. Recall tries to find all data points of interest and is assessed by how many items are correctly selected. These performances are measured using concepts known as true positive, false positive, false negative and true negative.

A true positive is when both the model and the reviewer agree on what the item is – for instance, that a red apple is a red apple. A false positive occurs when a model labels something as belonging that does not belong; for instance, labelling an orange as an apple. True negative is when the model and the reviewer agree that something does not belong in the class – for instance, the model finds an orange and says it is not an apple and the reviewer agrees that the label is correct, and that an orange is not an apple. A false negative is when the model says no apple is detected, when an apple is present.

Measuring precision and recall can help to identify whether more work needs to be done to improve the model (e.g. reducing false negatives or false positives), and model performance needs to address both precision and recall. In combination, these two measures are termed an F1-score, and it is the conventional approach to measuring model performance. F1-scores (as well as precision and recall, individually) produce values between 0 and 1.0, where scores of 0.8 and higher are considered good. The model performance details for this work are provided in Annex B.

Continuous feedback is needed for machine learning. This makes it both frustrating and tantalizing: good results can always be made better with more, higher-quality data.

Sources: **Porciello, J., Ivanina, M., Islam, M., Einarson, S. & Hirsh, H.** 2020. Accelerating evidence-informed decision-making for the Sustainable Development Goals using machine learning. *Nature Machine Intelligence*, 2(10), Article 10. https://doi.org/10.1038/s42256-020-00235-5

Porciello, J., Coggins, S., Mabaya, E. & Otunba-Payne, G. 2022. Digital agriculture services in low- and middle-income countries: A systematic scoping review. *Global Food Security*, 34: 100640. https://doi.org/10.1016/j.gfs.2022.100640

Understanding the data through document types

FAO plays a critical role in knowledge sharing and strategic thinking on issues that contribute to the transformation towards more efficient, inclusive, resilient and sustainable agrifood systems (FAO, 2021a). Much of this knowledge is shared with the broader community through processes of knowledge brokering, which include publication of numerous reports, books and more formal publications each year. Knowledge brokering comprises the largest subset of documents, nearly 45 percent of the dataset (Figure 7).

FAO also provides guidance to governments and other agencies on areas related to agrifood systems, and especially in areas where rapid growth is taking place, such as digitalization of agriculture, biotechnologies, food safety and renewable energy. The guidance is developed in accordance with member countries, and where input is captured through consultative processes, policy guidance reports and technical guidelines. These materials represent nearly 25 percent of the dataset. Technology utilized had the lowest representation, at just 10 percent of the dataset.

This distribution of 17 928 documents into the three categories (Figure 7) points to a slight gap in documents published by FAO in this dataset that highlight utilization of technology – whether it is through case studies, in-country assessments, or project and programme evaluation documents. Additional datasets curated by FAO, such as the FAO Digital Portfolio, could represent a useful supplement to this category.

FIGURE 7

Numbers of documents across knowledge categories: technical guidance, knowledge brokering, technology utilized. Definitions available in Figure 2

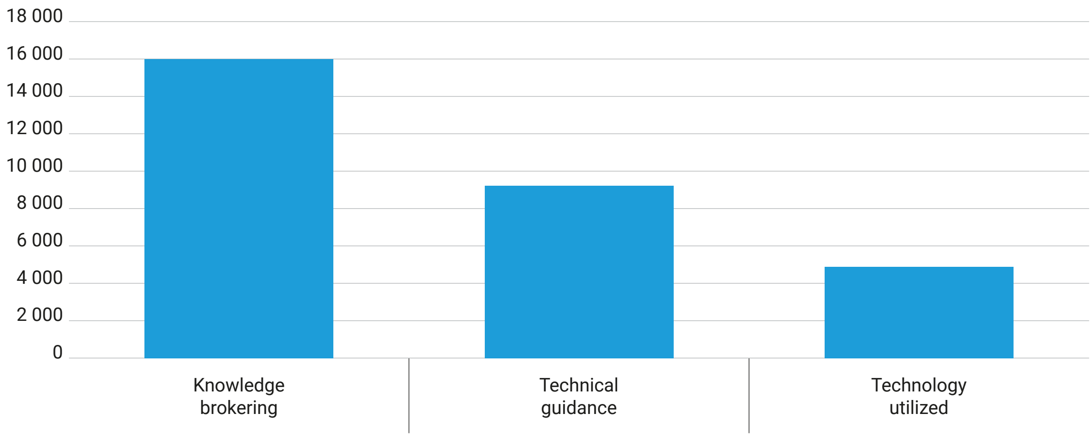
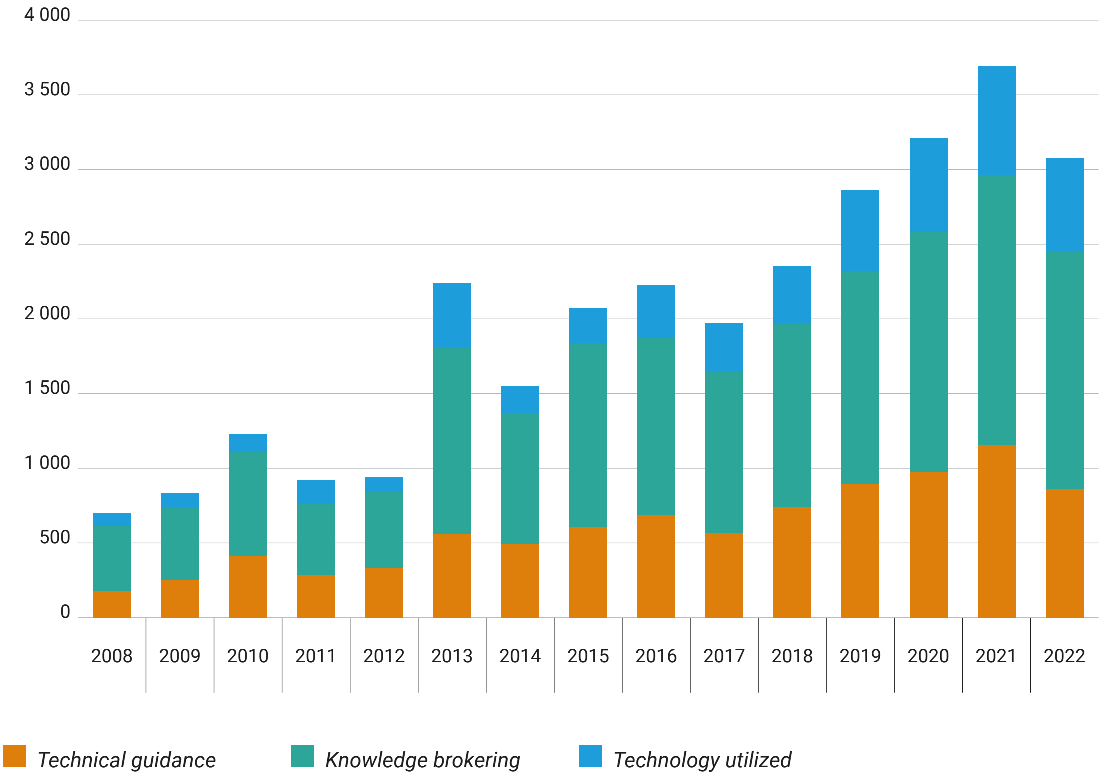


FIGURE 8

Annual distribution of documents across knowledge categories: technical guidance, knowledge brokering, technology utilized. Definitions available in Figure 2



Over the years, document types remained consistent across publications – with knowledge brokering continuing to represent the largest portion of documents each year. While the proportion of documents in each category remained relatively uniform across the years, in absolute terms there was growth in the number of publications that were either technology utilized or technical guidance and feasibility oriented (Figure 8).

Document category analysis provides a building block for the next phase of analysis to explore the density of technology types in this analysis (Glossary). For instance, if technologies were primarily mentioned in convening documents (such as conference notes or agendas), rather than in books or research reports, this could be viewed as a gap in FAO knowledge production and analysis that looks more in depth at technology. Alternatively, if a greater number of documents mentioning technology are manuals or training guides, then it is highly likely that FAO has been able to track technology implementation and utilization over the years.

Technology acceleration

The past decade has seen tremendous growth in agricultural technologies across low- and middle-income countries (FAO, 2022c; Reardon *et al.*, 2019). Given that the publication year parameter was 2008 onwards, it is consistent with an increase of more advanced technologies such as digital and biotechnological from 2008 onwards. A more detailed look at trends over the years as related to technology types can be found in Figure 9. Year-on-year growth for each technology type can be seen given that the number of accurately categorized documents increased with publication year. Incidentally, while the capture and analysis of "other" technologies was part of this exercise, they represented fewer than 30 documents in total.

FIGURE 9

Annual distribution of documents by technology type. Each document was classified with one technology category

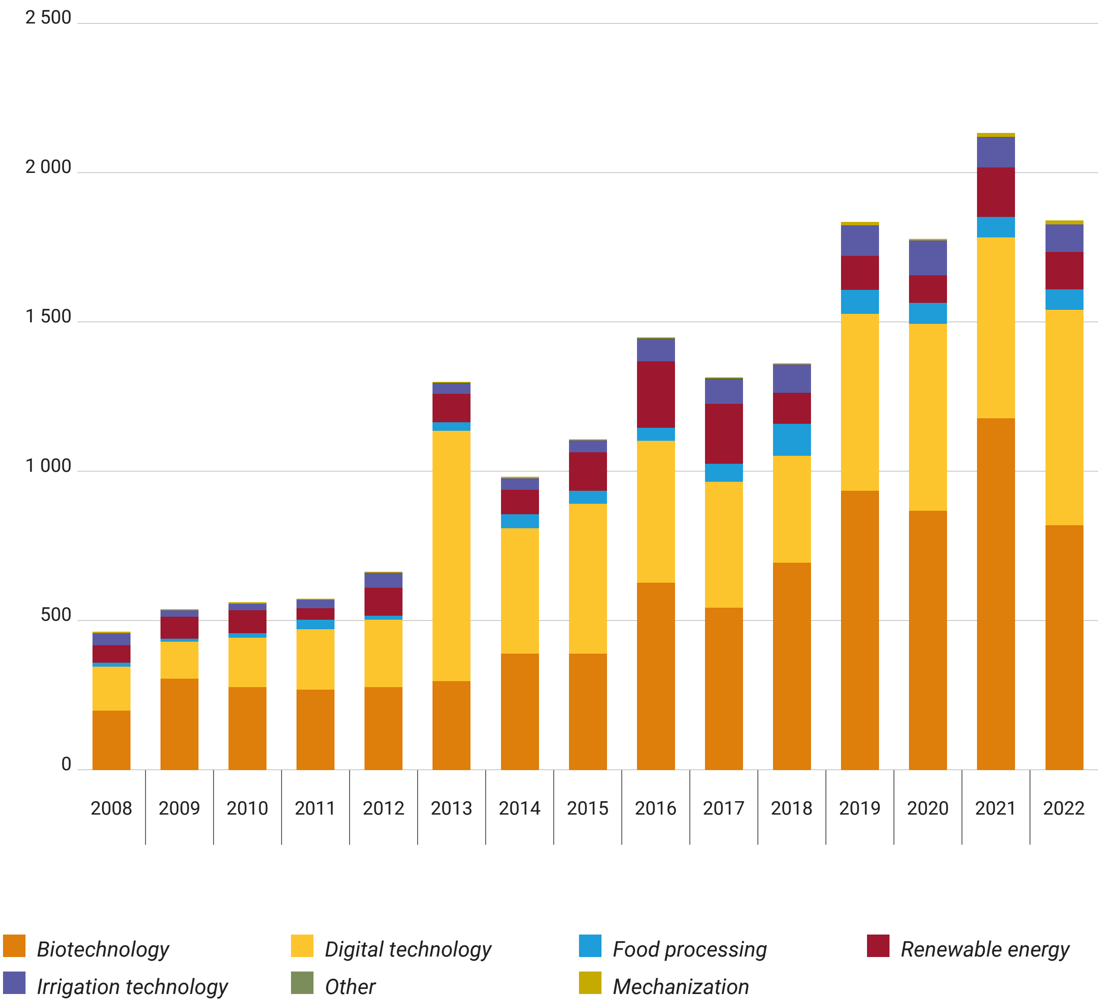
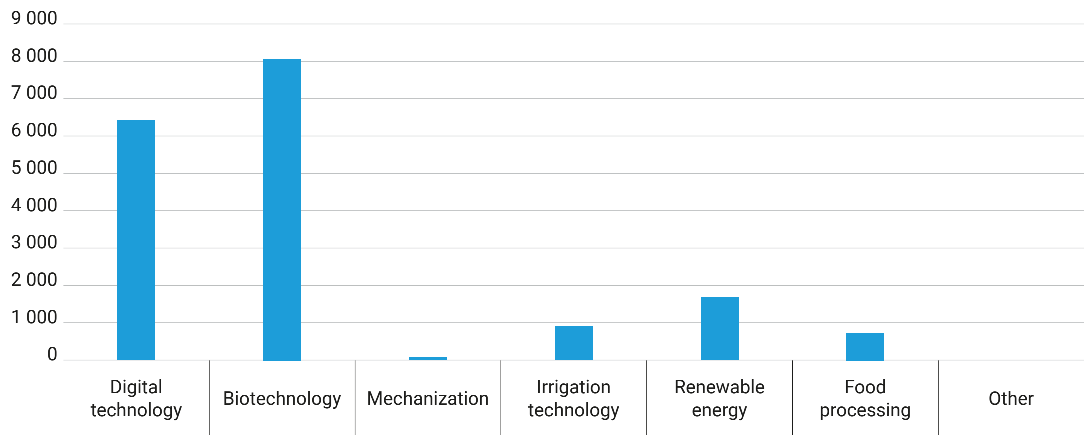


FIGURE 10

Numbers of documents categorized according to technology type



Numbers of documents by technology categories are given in Figure 10. Digital technology is a fundamental global driver, reshaping economies, governments and civil society – thereby impacting almost every aspect of development. The literature on digital technologies for agriculture has focused on potential use of advanced technologies such as drones, sensors, and AI, much of which require upgrades to infrastructure before improvements to agricultural outcomes for small-scale family farms can be realized (Chandra and Collis, 2021). Such upgrades are similarly required to improve digital services for farmers (Fabregas *et al*., 2019).

Digital technologies are a necessarily broad category, but some trends stand out. The opportunities for digital technology to provide support to the entire agricultural sector to be more efficient and interconnected, especially through private sector engagement, is rapidly emerging as an important topic: in the three-year period from 2019 to 2022, nearly 40 documents highlight the private sector as a central element in digital technology, as compared with only 20 documents in the previous ten years combined. The FAO Strategy for Private Sector Engagement 2021–2025 (FAO, 2021b) emphasizes the importance of the private sector to bring about transformative change.

Nearly 20 percent of all the biotechnology documents focus on topics related to pesticide use, compliance and enforcement, and guidelines and assessment, representing an emphasis on FAO guidance and knowledge needed to support a shift in the world's production practices that encourage reduced synthetic inputs in favour of mechanisms that contribute to planetary health. The focus on use and impact of pesticides in pest and plant management control includes offering scientific advice to address the issue of unintended presence of residues in crops, animal feed and veterinary drugs and medicine, representing a wide spectrum of topics and needs for both governments and practitioners, including extension and advisory agents.

Traditionally, FAO has provided consistent global guidance on biotechnologies through consultative processes. A recent report highlights that across low-and middle-income countries, technologies such as tissue culture, marker-assisted selection and artificial insemination are still used extensively and with good results due to their wide availability, low costs and existing infrastructure (FAO, 2022d).

As biotechnology continues to evolve, additional guidance will be needed for Members. Currently, very few publications focus on emerging technological innovations related to biotechnology, such as CRISPR (clustered regularly interspaced short palindromic repeats) and other gene-editing techniques (FAO, 2022b),[3] synthetic biology and cellular agriculture. Some documents focus on the role of gene-editing in aquaculture. An increase in the knowledge, guidance and feasibility for this information may contribute towards greater understanding of the use of these technologies and their potential application, including in climate change adaptation and mitigation.

Food processing technologies have created new markets for producers, along with employment in various supply-chain segments, including food processors, wholesalers and logistics firms and enabling small-scale producers (Reardon *et al*., 2019). They rely on the value chain to sell their products, receive logistics and intermediation services, and buy farm inputs.

Using technologies and methods to gain efficiencies across the system by reducing food loss and waste, and improving food safety, is a prevalent theme in the dataset. Yet gaps in the knowledge associated with food processing, including food processing technologies, have been identified in other studies and are similarly mirrored in the FAO analysis. Research has referred to the supply chain segments that provide food processing technologies as a "hidden middle", even though they constitute a significant percentage of actors in an average food supply chain (Reardon *et al*., 2019). A recent scoping review reported a dearth of empirical evidence "on the role that SMEs in the midstream and downstream of input and output value chains can play in the adoption and dissemination of agricultural practices that will preserve the environment or increase small-scale producers' resilience to climate change" (Liverpool-Tasie *et al*., 2020). Increased knowledge and awareness about scaling food processing technologies at the subnational level can promote opportunities to reduce post-harvest food loss and waste and increase off-farm economic growth year-round for communities. It may also provide insight into increased opportunities to improve dietary diversity if such technologies emphasize processing a diverse set of underutilized and nutritious crops that reflect local preferences.

Mechanization is a key contributing factor in discussions that are focused on productivity, employment and transformation of agrifood systems. The FAO State of Food and Agriculture 2022 emphasizes agricultural automation as essential to plugging the labour gap, given that many manual tasks associated with fruit and vegetable production are low paid and represent difficult working environments.

Discussions about mechanization fall into two primary thematic categories: ensuring the mechanization frameworks are linked to value chains, investment frameworks and services for rural communities, and more technical and practical documents focused on training on equipment and other documents. In addition, cross-cutting themes such as youth and child labour are discussed by Takeshima and Vos (2022), and gender issues are evident in other FAO analyses on mechanization (Justice *et al*., 2022).

3 FAO published an issue paper on gene-editing techniques and agrifood systems in December 2022 that presents a balanced discussion of the most pertinent aspects of gene editing, including the consequences for human hunger, human health, food safety, effects on the environment, animal welfare, socioeconomic impact and distribution of benefits.

Across all categories, mechanization was associated with the least number of documents. A gap in materials focused on mechanization is the result of some crossover with technologies that are classified under digital and irrigation technologies. For instance, agricultural robots are one of the keywords for mechanization, yet robots and their role in supporting precision agriculture is a frequently cited example of digital agriculture, and therefore would crossover into digital technologies. Similarly, while solar-powered irrigation pumps contribute to on-farm mechanization and productivity, such mechanical technologies are captured under irrigation, not mechanization.

In the context of agrifood systems, the integration of diverse technologies across different categories is essential to create improved products that benefit all producers, both on- and off-farm. On the one hand, there is potential risk that classifying knowledge products into categories may inadvertently minimize technology readiness, transfer and leapfrogging across different domains. After all, an ensemble of different technologies is needed to power farms, markets and an enabling infrastructure. For instance, remote sensing technologies are a common feature of modern tractors that perform precision agriculture functions. On the other hand, results such as these can be instrumental in ensuring that historical concepts and definitions are periodically updated and remain representative.

The water – energy nexus continues to be a major topic of discussion as demand for energy rises and more countries become water constrained. Finding ways to improve irrigation practices by making them less energy intensive has been a focus for recent technology innovations (FAO, 2022e). Renewable energy can be a reliable alternative to address increasing demand for energy and water resources, whether it is through solar pumping, desalination, or reducing water usage across other intensive sectors (such as shifting away from fossil fuels to more renewable sources). Solar pumping and solar-powered desalination plants can decrease costs and increase production capacity, while being sustainable ways to approach growing water constraints. Although renewable energy and irrigation were only a small portion of technologies identified in this dataset, it is likely that in the coming years there will be more research at the intersection of these categories as more investment is put into green technologies to address water scarcity.

Examples of categorization

Accurate high-quality data analysis is challenging, especially as organizations begin to use more sophisticated approaches to manage data and knowledge, including an increased reliance on dashboards and data analytics. Most organizations will require a multi-pronged approach. A persistent question remains about how knowledge products that cover the breadth and depth of complex, current questions are accurately captured, and where tensions among categories might occur.

This section provides a detailed look at four documents that were categorized. These were selected to demonstrate the complexity of the process, and importantly, the role that expert user feedback plays to continuously update and help fine-tune machine-modelled approaches. Boxes 3 and 4 highlight the breadth and depth of research as well as the challenges in analysis.

BOX 3

Streamlined analysis with clear results

In some cases, the analysis was able to accurately identify documents that fit a certain document and technology category based on the metadata available.

Document 1: Forestry for a low-carbon future: integrating forests and wood products in climate change strategies (FAO, 2016)

This document is an example of **knowledge brokering** on **renewable energy** technologies.

Abstract: "Following the introduction, Chapter 2 provides an overview of mitigation in the forest sector, addressing the handling of forests under UNFCCC. Chapters 3 to 5 focus on forest-based mitigation options: afforestation, reforestation, REDD+ and forest management, and Chapters 6 and 7 focus on wood-product based options, wood energy and green building and furnishing. The publication describes these activities in the context of UNFCCC rules, assessing their mitigation potential and economic attractiveness as well as opportunities and challenges for implementation. Chapter 8 discusses the different considerations involved in choosing the right mix of options as well as some of the instruments and means for implementation. Chapter 8 also highlights the co-benefits generated by forest-based mitigation and emphasizes that economic assessment of mitigation options needs to take these benefits into account. The concluding chapter assesses national commitments under UNFCCC involving forest mitigation and summarizes the challenges and opportunities."

Document 2: Kenya irrigation market brief (FAO/IFC, 2015)

This document is an example of **knowledge brokering** on specific **irrigation** technologies.

Abstract: "Achieving Africa's agricultural growth potential will require a significant increase in historically low levels of productivity. This is an area where irrigation can play a critical role. Modern, efficient irrigation systems can substantially increase crop yields, resulting in improved livelihoods, reduced risk associated with drought, efficient use of limited water resources, and greater food production. This report is the fourth in a series of market briefs produced jointly by IFC and the Food and Agriculture Organization of the United Nations (FAO). It is targeted primarily at private sector investors and companies interested in expanding investment in irrigation in sub-Saharan Africa, with particular focus on modern irrigation technologies, but may be of wider interest to all stakeholders engaged in irrigation development in the country. The report assesses the current state of the irrigation market in Kenya, recent performance, and opportunities for future growth."

Sources: **FAO**. 2016. Forestry for a low-carbon future: Integrating forests and wood products in climate change strategies. Rome, FAO. https://www.fao.org/3/i5857e/i5857e.pdf

FAO/IFC. 2015. Kenya: irrigation market brief. Rome, FAO/Washington D.C., IFC. https://www.fao.org/3/i5074e/i5074e.pdf

BOX 4

Examples of text summaries that provide insight into the challenges of hybrid human-machine text classification

While some classifications were easy to identify and validate, others were not as easily analysed by the model and presented validation challenges for FAO.

Document 3: Water accounting and auditing: A sourcebook (FAO, 2016)

This document is an example of **technology guidance** and feasibility based on current definitions. It provides practical advice and informs guidelines for the water sector. But while the document is about water management, it does not specify irrigation technologies as defined in the Glossary. There are indications that this document might include approaches to digital technology, based on language around monitoring hydrological cycles and a transition to a sustainable agricultural production system.

Abstract: "The rationale behind these water accounting and auditing guidelines is that scope exists worldwide to improve water-related sectoral and inter-sectoral decision-making at local, regional and national levels. In many regions of the world, sustainable and reliable delivery of water services has become increasingly complex and problematic. Complexities that are very likely to increase, considering the unprecedented confluence of pressures linked to demographic, economic, dietary trends, and climate change. Particularly if overall demand for freshwater exceeds supply, the delivery of water services is often less about engineering, although engineering is still required, and more about politics, governance, managing and protecting sources, resolving conflicts about water, ensuring rights to water are respected, and so on. It is also about understanding and monitoring the hydrological cycle at the appropriate scale of analysis. This is where water accounting and auditing can play a crucial role. The rationale behind this water accounting and auditing sourcebook is that scope exists worldwide to improve water-related sectoral and inter-sectoral decision-making at local, regional and national levels. Water accounting and auditing are recommended by FAO and others as being fundamental to initiatives that aim to cope with water scarcity. This sourcebook aims to provide practical advice on the application and use of water accounting and auditing, helping users planning and implementing processes that best fit their needs."

Document 4: Business models along the poultry value chain in Kenya: evidence from Kiambu and Nairobi City Counties (FAO, 2022)

The **food processing technologies** category applied to this document is relatively clear. The document provides guidance on business models for small-and-medium sized enterprises in the poultry value chain. Yet, "biosecurity practices" was captured as a possible technology intervention, which could have resulted in classifying this document as "biotechnology". Classifying it as "biotechnology" and ignoring its applicability to food processing would not be intuitive.

Abstract: "FAO is supporting a One Health dialogue in Kenya to facilitate the adoption of biosecurity practices along the poultry value chain. This study characterizes the business model and the enterprise budget of farmers, traders, processors and retailers along the poultry value chain in Kiambu and Nairobi City Country. Results show that all poultry businesses are profitable and avail resources to start adopting biosecurity practices that minimize the introduction and spread of pathogens in animals. However, while producers have some incentives to adopt biosecurity practices, as avoiding and controlling diseases is essential for their profitability, traders, processors and retailers have little if any incentives to adopt biosecurity practices as they keep birds or poultry meat only for a limited period of time and the market for poultry products do not differentiate between safe and unsafe products. It is essential that animal health services systematically provide services beyond the farm gate to minimize the public health risks along the poultry value chain."

Sources: **FAO**. 2016. Water accounting and auditing: A sourcebook. Rome, FAO. https://www.fao.org/3/i5923e/i5923e.pdf

FAO. 2022. Africa Sustainable Livestock 2050: Business models along the poultry value chain in Kenya – Evidence from Kiambu and Nairobi City Counties. Rome, FAO. https://doi.org/10.4060/cb8190en

Examining outcomes mentioned

An outcome is typically the effect or the result of an activity, programme, strategy or intervention. Outcomes play an important role in the assessment of both success or failure of a programme, and assessment of an outcome involves a carefully planned methodology and framework (Munn *et al*., 2018).

Agricultural research has been traditionally focused on improving productivity of a small number of staple crops (Pingali, 2015). There is a major shift in thinking about agriculture, one which puts agriculture in the larger context of an agrifood system with complex interactions among nutrition targets, gender and social inclusion of marginalized groups, biodiversity and climate change (Lipper *et al*., 2020). There are positive and negative interactions that occur across agrifood systems, leading to both intended and unintended consequences. For instance, a lack of research on fruits, vegetables and nutritious grains like millet and sorghum, as well as accompanying post-harvest storage to ensure safety and reduced loss, are examples, relative to the major crops, of where there are significant risks for poor diets and subsequently micronutrient deficiencies and malnutrition (Stathers *et al*., 2020).

Evaluating the outcomes mentioned in a dataset such as this is one way for analysis to be more impact driven. Technology should not be isolated from its impact on users and intended and unintended communities. A bird's-eye view of outcomes mentioned is distinct, however, from measuring and assessing outcomes in exercises such as impact evaluations and evidence syntheses, including systematic reviews. Additional analysis is needed across all documents to determine whether outcomes were mentioned in relation to the technology, or whether an assessment of a series of outcomes took place within the report. However, the value of identifying outcome trends, as conducted in this pilot analysis, indicates areas of interest for the agricultural development community at any point in time, and potentially aspirations as to where it might be heading to in the future.

Figure 11 shows the volume and categories of outcomes mentioned in 10 463 documents with title and summary data, using a multi-label model (outcome definitions in Annex B). The model is trained to assess the concept of an outcome – as opposed to relying only on keywords – before assigning the outcome to a category. Outcomes that are detected, but do not map to a category, are also captured and included in the "other" outcome category.

Each document references, on average, two outcomes. Most of the outcomes mentioned across the documents support an increase in food security and nutrition, followed by outcomes focused on incomes, yields and productivity.

Table 1 identifies the links between mentioned outcomes and technology types across all documents. Outcomes related to economic growth, such as incomes and productivity, and food security and nutrition, are most mentioned across all publications, but areas that are high priority, including environmental outcomes and social inclusion, are some of the least mentioned. Biotechnologies can lead to improvements in crop varieties and livestock breeds to increase their productivity as well as enhance the nutritional value of major food staples (FAO, 2022d). Similarly, digital agriculture, and in particular precision agriculture, can improve the livelihoods of farmers through increases in income, productivity, employment and practice change.

The area with the least mentions encompasses environmental outcomes. Given the well-documented negative impacts of agriculture on the environment, it is concerning that outcomes related to climate and environmental sustainability across all technologies are low. There is a dearth of data about how digital agricultural services can be used by farmers to adapt to a changing climate, where some studies have failed to incorporate data that would have demonstrated improved environmental or climate links, such as measuring water conservation, reduced leaching of nutrients into the environment, or use of fossil-fuel-reliant irrigation pumps (Porciello *et al*., 2022; Ricciardi *et al*., 2020).

FIGURE 11

Frequency of outcome categories across 10 436 studies that included both title and summary data

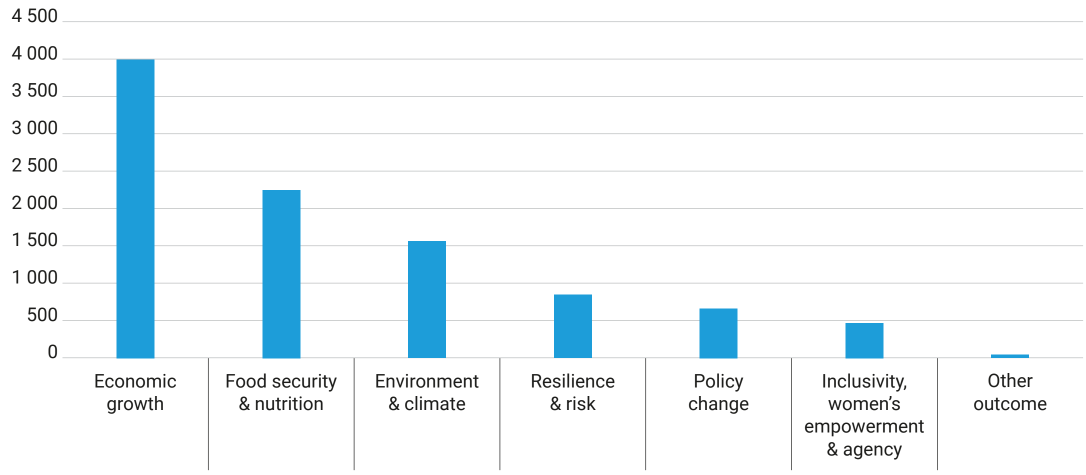


TABLE 1

The relationship between outcomes mentioned and technology types identified, based on the number of documents

	Biotechnology	Digital technology	Irrigation technology	Renewable energy	Food processing	Mechanization	Other
Economic growth	1 292	1 168	260	261	214	24	4
Environment & climate	508	358	136	193	49	4	2
Food security & nutrition	923	593	96	67	151	5	1
Inclusivity, women's empowerment & agency	168	139	22	17	18	2	1
Policy change	328	147	31	30	21	3	0
Resilience & risk	362	185	41	41	24	1	0

Social and demographic details

FAO continues to pay special attention to the specific contexts of the communities it works with to ensure that the introduction of a new technology does not unintentionally perpetuate or create new inequities. While the use of a new digital technology may theoretically allow for maximum efficiency, such technologies, which are typically developed in wealthier, high-income countries, could unintentionally exclude the nearly 3.8 billion people that still lack access to the internet, whether due to broadband issues, lack of a mobile device or other barriers (GSMA, 2022).

Better identification of relevant characteristics of people, both individuals and communities involved in agricultural activities, is essential. To answer specific questions about which people are impacted by agrifood systems technologies, data about specific populations that interact with and are afforded use of various technologies must be considered an essential component of research and development. Higher quality data collection practices are becoming more commonplace and standardized in many nationally representative surveys, especially as the uptake of measurement instruments, including the Women's Empowerment in Agriculture Index (WEAI), continues to increase (FAO, 2023).

However, a lack of data and underreported data on social and demographic details about specific populations outside nationally representative surveys, including women, is a major impediment to understanding opportunities about how to make agricultural technology more equitable. A scoping review on digital agriculture technology services for farmers reported that fewer than 30 percent of all research papers evaluated included data about study populations and their sociodemographic factors (Porciello *et al.*, 2022). A 2020 editorial from Nature highlighted this issue, stating that only 2–3 percent of research captured the population of study across eight systematic scoping reviews (Nature, 2020).

Table 2 shows the classification schema for how populations in this dataset were identified. It is noted that there are additional communities and populations that could have been identified, and have been successfully identified in other exercises, including Indigenous Peoples, vocation types, specific ages or phases of life, such as nursing mothers or pregnant women. This report begins to explore this approach of identifying populations that interact with technology using a limited number of high-level categories.

TABLE 2

An overview of the basic demographic descriptions, generated through descriptive data from text-based studies, such as reference to the age or sex of a study population

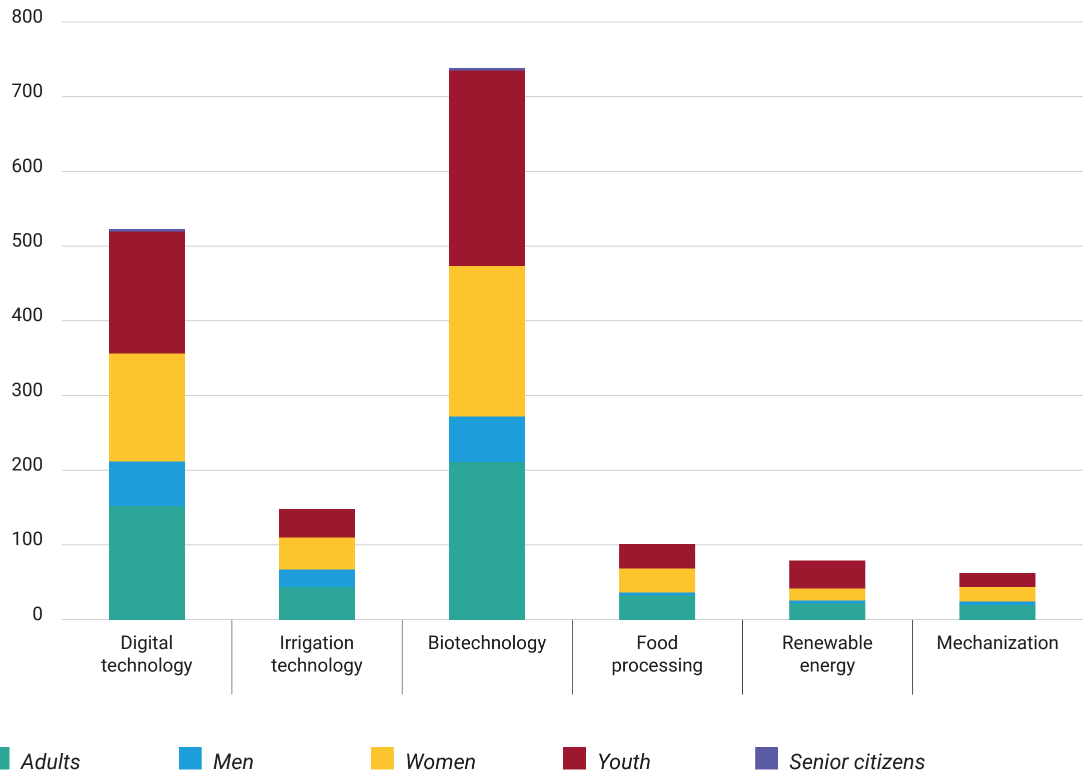
Study populations referenced in Figure 12	Study population descriptions associated with the primary label (not exhaustive)
Senior citizens (sex-disaggregated)	Senior citizen, senior person, elderly, pensioner, retiree, grandmother, grandfather
Adults	Smallholder farmer, farmer, small-scale producer, producer, grower, family farmer, adults, grown-up, widow, (without specification) head of household, farm, land, or home
Men	Man, men, male, husband, father. Also, smallholder farmer, farmer, small-scale producer, producer, grower, family farmer, with any of man, men, male, husband, male/man head of household, farm, land, or home
Women	Woman, women, female, wife, mother. Also, smallholder farmer, farmer, small-scale producer, producer, grower, family farmer, with any of woman, women, female, wife, mother, head of household, farm, land, or home
Youth (not sex-disaggregated)	Youth, young people, children, teenager, boy, girl, son, daughter. Also, school, school-age, pupil, learner

These details, despite trying to capture all variations, were provided in extremely low numbers across the data. Figure 12 identifies user groups per technology type and shows the number of documents mentioning populations, for instance, women were identified in 648 documents. Such low reporting, while not uncommon, makes it impractical to draw clear conclusions about how any user group is meaningfully engaging and using any technology across all categories. There is a demonstrable lack of research on senior citizens, independent of the technology type being addressed in the research.

Overlapping factors such as age, sex, class, race, and socioeconomic status, can create intersectional and interdependent systems of discrimination and disadvantage which reinforce the exclusion of some groups from the benefits of technology. The ability to examine technology adoption, uptake, and use across multiple sociodemographic dimensions is essential to foster greater understanding of technology acceleration and use and to avoid unintentional bias or exclusion (see Box 5).

FIGURE 12

Frequency of social and demographic details mentioned according to technology type



BOX 5

Opportunities for social factor data capture

In addition to leveraging the power of AI to capture qualitative data, simple tools like input forms could identify populations and communities of study, outcomes described, average sample sizes of population, methodologies used in assessment, and would enhance the analytical capacity of descriptive data coming from FAO materials.

There are some publications that lead the way in providing high-quality and visible information about study populations. Such examples illuminate how simple questions prior to the publication of any FAO document, and using controlled vocabulary responses and displaying in an interactive dashboard, would provide an innovative way of helping users engage with FAO data, especially on cross-thematic issues like populations and outcomes.

A 2017 FAO publication addressed the issue of social protection to foster sustainable management of natural resources and reduce poverty in forestry-dependent communities (FAO, 2017). The qualitative approach sought to determine and analyse the diversity of social protection needs and opportunities for forest-dependent communities in five districts of Uganda and make recommendations for addressing them. Sampling covered 29 villages and 322 households, 12 focus group discussions involved 229 men, 117 women and 23 young people, and 41 interviews with informants were carried out at the district level, with six at the national level.

Source: **FAO**. 2017. A mapping of social protection needs and opportunities for forest-dependent communities in Uganda. Rome, FAO. https://www.fao.org/3/i8061e/i8061e.pdf

Conclusions

The emergence of new technologies across the agrifood sector is a signal of transformation and a period of rapid economic growth. Evaluating the evidence base about specific technologies is an important approach that helps make better data-driven assessments about future growth and identify where gaps in the data might exist.

Some of the gaps in the dataset reflect overlaps between the technology type categories that may not be indicative of widespread and pervasive gaps. For instance, mechanization and irrigation exhibit some overlap with the descriptions and examples of digital technology, potentially leading to identification of gaps in coverage that are less about actual gaps and more about lack of clarity within the definitions themselves. In addition, categories such as digital technologies are by default much broader and encompass more narrow categories, leading to an information asymmetry issue.

While an additional expert review of the technology type categories would ensure that the distinctions are clear enough for future assessment, robust coordination that goes beyond stocktaking activities is needed.

This report highlights a significant gap in data collection practices and these gaps are indicative of more structural issues regarding concepts and definitions in the agrifood sector. In comparison with sectors like health and medicine, the agrifood sector lacks analytical frameworks, ontologies, and taxonomies. For example, the World Health Organization (WHO) has established the International Classification of Health Interventions, whereas the primary structured collection of agricultural concepts in agrifood systems, the FAO multilingual thesaurus, merely defines an intervention as a "controlled price". Similarly, the controlled thesaurus maintained by the United States National Agricultural Library (NAL) does not encompass agricultural intervention concepts or definitions.

Advancing opportunities for assessment is important for all sectors, but perhaps no more so than for food security and agriculture at this moment in time. FAO's consultative and administrative capacity represents an advantage over other organizations to produce thoughtful leadership on this issue. FAO's new global knowledge product, ATIO, aims to curate existing information on the current, measurable state of science, technology and innovation and upcoming changes, as well as their transformative potential, to inform evidence-based policy dialogue and decision-making, including those on investments (FAO, 2022f).

Despite increased attention by governments, funders and the research community to inclusivity and other metrics linked with the well-being of small-scale producers, such as resilience, there is no standard assessment for technology use among communities. Such data are only sporadically captured or reported.

As discussed in this report, technologies are not inherently good or bad, but the promotion of new technologies in poorly resourced communities requires nuance and knowledge about the communities. On the one hand, such communities may either suffer from a lack of infrastructure or may even frequently deal with natural hazard-induced disasters that deny any opportunities to establish permanent infrastructure. On the other hand, countries that are more technologically advanced may be able to adopt and experiment with technologies in ways that benefit them and their communities. Being able to assess which groups can access certain technologies, as well as facilitators and barriers, is critical for deploying technologies more equitably.

FAO has been studying technology use for decades. Better intelligence and insight on how technologies are impacting communities requires data capture on community use and outcomes, including unintended consequences. Additional dialogues with user communities to establish how the technology is being adopted, and whether new and novel uses are taking place, is also needed. Currently, the gaps in contextual understanding of technology use in communities as presented across a range of different knowledge products, represents a challenge to understanding the intended meaning, potentially resulting in misinterpretation or inaccurate analysis. Insufficient knowledge can impede the effective application of information for learning purposes.

Recommendations

Further review technology type categories

Some of the gaps in the dataset reflect overlaps between the technology type categories that may not be indicative of widespread and pervasive gaps. For instance, mechanization and irrigation exhibit some overlap with the descriptions and examples of digital technology, potentially leading to identification of gaps in coverage that are less about actual gaps and more about lack of clarity within the definitions themselves. In addition, categories such as digital technologies are by default much broader and encompass more narrow categories, leading to information asymmetry. An additional expert review of the technology type categories would ensure that the distinctions are clear enough for future assessment and stocktaking activities so that identified gaps reflect important gaps.

Enhanced data about communities must be prioritized

The dearth of data concerning social and demographic communities represents an urgent and critical gap, limiting FAO's ability to connect social inclusiveness to social impact. Leveraging AI to conduct additional analysis across the full text represents one opportunity to fill the gap, but retrospective analysis will not change normative behaviour across the research and policy communities to encourage better data collection practices. More work is needed to support working groups that help to address the issues of data collection at scale. Encouraging better practices can include processes that encourage standardizing data analysis, such as creating standard input fields and forms for future FAO documents. Such techniques would generate a wealth of information for analysis and would result in greater opportunities for use/uptake/understanding, including making annual reporting for internal use more dynamic and impactful.

Integration of trend and investment data to strengthen impact

Knowing what portion of FAO's investment in technology aligns with the trends in this evaluation will contribute towards consistent, ongoing evaluation in the future. An analysis is needed at the activity and output level across FAO projects to identify how technology is used, and which facilitators and barriers exist across communities.

Generate demand for FAO data through new interfaces informed by consultative processes

The capture and use of quantitative data as analytics provides more interesting and useful ways for FAO to summarize the rich findings of its collection. New interfaces, including dashboards, that allow users to drive their own research and exploration through similar AI and expanded approaches used in this analysis offer rewarding opportunities for both internal and external FAO audiences to find and explore content. There are promising approaches delivering on AI-generated metadata that could significantly boost FAO's ability to produce baseline assessments. However, feedback from the community, through upfront and iterative on-going processes, is essential for any new system and approach to work. While persons familiar with the programming activities across geographies and programme areas will play a crucial role in generating inputs needed for the analysis, and in reviewing and validating the model results, these persons should not guide the quest for new systems. To generate demand for the uptake and use of the abundant knowledge that has been produced, there must be a user-driven process from the outset. FAO's strength in the consultative process for technical areas must be expanded to consider design of up-to-date systems.

Invest in knowledge management systems and new skills as part of the agrifood systems transformation

The combination of multiple legacy systems and an absence of metadata impacts document extraction and analysis, particularly when thousands of documents are available. The lack of available summaries across FAO metadata is just one concern for rapid, future data analysis.

New data science approaches require planning, investment and opportunities for testing and validating the approaches. For projects such as this one, the next stage of evaluation for transforming knowledge management approaches is to pilot their use, and support ongoing reports and decision-making panels to gauge whether improved analytics made available through ML categorization and classification are beneficial to research and implementation teams. Given the rapid expansion of growth in the sector that reflects the ever-changing needs of users to access and make use of FAO knowledge, it requires the Organization to be flexible regarding testing and establishing new knowledge management approaches.

Annex A
Technology type keywords

Technology	Keywords (FAO)	Extended synonym search
DIGITAL TECHNOLOGY	• App • Artificial intelligence (AI) • Big data • Blockchain • Cloud service • Cloud services • Crowdsourcing • Cybersecurity • Digital advisory (automated-voice response) • Digital advisory (smart phone/tablet) • Digital advisory (SMS) • Digital finance • Distributed Ledger Technologies (DLT) • Drones • E-commerce • eLocust3 • FAMEWS • FAOSTAT • Geographic information system (GIS) • Geospatial • Hand-in-Hand Geospatial Platform • Information and communications technology (ICT) • Internet of things (IOT) • Machine learning (ML) • Mobile data collection • Open Foris • Precision agriculture • Precision fish farming • Remote sensing satellite imagery • SEPAL • Virtual/augmented reality • WaPOR	• arcgis • augmented reality • cloud computing • digital agriculture • digital technology • digital twin • geomatics • georef- • global positioning system • GPS • mobile device • mobile telephone • precision farming • QGIS • remote sens* • robotics • satellite navigation system • smart agriculture • smart device • smart sensor • smartphone apps • smartphone* • speech recognition • virtual reality • wireless sensor network

Technology	Keywords (FAO)	Extended synonym search
BIOTECHNOLOGY	• Alternative protein sources • Animal genetics/livestock genetics • Animal/Livestock vaccines • Artificial insemination • Bioaugmentation • Biocontrol • Bioeconomy • Biofertilizer • Biofortification • Bioinformatics • Biopesticide • Bioremediation • Bioremediation • Biosafety • Biostimulant • Cell-based meat • Cellular agriculture • Chromosome set manipulation • CRISPR-Cas • Cryopreservation • Digital sequence information (DSI) • DNA barcoding • DNA sequencing • Enzyme-linked immunosorbent assay (ELISA) • Estrus synchronization • Fermentation* (mentioned in food processing) • Gene editing • Gene sequencing • Genetic modification (GM) • Genetically modified food • Genetically modified organism (GMO) • Genome editing • Genomic selection • Genomics • Green plant protection • In vitro fertilization • In vitro slow growth storage • Lab-grown dairy • Living modified organism • Microbiome • Micropropagation • Molecular marker-assisted selection • Molecular markers • Mutagenesis (chemical, physical) • Mutant varieties • Mutation breeding • Next Generation Sequencing • Nuclear technique/technology • Omics • Plant genetics/Crop genetics • Polymerase Chain Reaction • Precision fermentation • Probiotic • Progesterone monitoring • Radioimmunoassay • Reproductive biotechnologies • Sexing • Somatic hybridization • Sterile insect technique (SIT) • Synthetic biology • Transgenesis/transgenic • Whole genome sequencing • Wide crossing	• biomedicine • biopharmaceutical • biopharmaceuticals • bioscience • functional genomics • genetic manipulation • genetic resource • genetic transformation • metabolic engineering • molecular biology • nanobiotechnology • nutrigenomics • nutritional genomics • recombinant DNA technology

Technology	Keywords (FAO)	Extended synonym search
MECHANIZATION	• Agricultural Robots • Bailers • Combine harvester • Direct Seeding and CA equipment • Drones for agriculture application • Fertilizer application machines • Fisheries equipment • Forage Harvester • Forestry equipment • Hand hoe • Harvester/ harvesting tool • Irrigation equipment • Land clearing equipment • Land Preparation Equipment • Livestock equipment • Milling operations tool • Mini motor cultivator • Pesticide application machines • Pickaxe • Processing equipment • Rake • Repair and maintenance tool • Seeding and planting tool • Shovel • Solar dryer • Solar energy equipment • Storage equipment • Thermal dryer • Threshing dehulling equipment • Tractors (two-wheel tractor / four-wheel tractor) • Transportation equipment • Watering can • Weed control tool • Wheelbarrow	• agricultural machinery • automation • commercialization • drudgery reduction • extension service • harvesting machine • industrialization • mechanical harvesting • mechanized operation • mechanized planting • modernization • power tiller • shoring

Technology	Keywords (FAO)	Extended synonym search
IRRIGATION TECHNOLOGY	• Basin development • Basin irrigation • Californian network distribution system • Corrugation irrigation • Deficit irrigation • Desalination for irrigation • Drip irrigation • Furrow irrigation • Gravity irrigation • Inland valley bottoms • Irrigation efficiency • Irrigation infrastructure • Irrigation modernization • Localized irrigation • Manual water extraction technologies (Pumps and canals) • Mechanized water extraction technologies (heat and motor pumps) • Mechanized water harvesting technologies boreholes and other wells (heat and motor pumps) • Power irrigation • Solar irrigation • Solar irrigation • Spate irrigation • Sprinkler irrigation • Supplementary irrigation • Tele-irrigation • Tidal irrigation • Use of water • Wastewater irrigation • Water harvest • Water infrastructure • Water management	• centre pivot irrigation • drip irrigation • infrastructure • irrigation scheme • laser land levelling • lift irrigation • micro-irrigation • motor pump • mulched drip irrigation • power tiller • pressurized irrigation system • solar pump • subsurface drip irrigation • treadle pump • trickle irrigation • tubewell

Technology	Keywords (FAO)	Extended synonym search
RENEWABLE ENERGY	• Solar • Bioenergy • Biofuel • Biogas • Clean cookstoves • Hydropower • Wind • Wood energy	• bio ethanol • biodiesel • bioeconomy • bioethanol • biofuels • bioheat • biomass energy • biomethane • biorefineries • biorefinery • clean energy • cogeneration • desalination • diesel generator • direct combustion • electric vehicle • electricity generating • electricity generation • electricity mix • fuel-efficient cooking stoves • geothermal • geothermal energy • geothermal heat • grid electricity • hydroelectricity • hydrothermal • mini grid • naturally replenishing energy • ocean energy • photovoltaic • power generation • powered renewable energy • renewable • renewable electricity • renewable energy • renewable fuel • renewable source • rural electrification • seawater desalination • solar energy • solar photovoltaic • solar photovoltaics • solar power • solar PV • solar water • solar wind • sunlight energy • sustainable energy • wind energy • wind power

Technology	Keywords (FAO)	Extended synonym search
FOOD PROCESSING	• Active/Smart packaging • Adding chemical preservatives such as sodium metabisulphite or sodium benzoate • Adding citric acid or vinegar • Canning • Centrifugation • Chilling • Cold chain • Concentrating by boiling, filtering • Cooling • Design, management and operation of fresh produce Packinghouse • Drying • Extraction of oils (ex. virgin coconut oil) • Fermentation* (mentioned in biotechnology also) • Freeze drying • Freezing • Frying • Grains and pulses, metal silos, hermetic bags • Heat pump drying • Hot water treatment for fruits • Hurdle technology • Increasing acidity • Juice extraction • Microfiltration of coconut water • Modified atmospheric storage and packaging • Non-thermal processing • Packaging • Pasteurization • Pickling • Post-harvest technology • Pressing • Processing of fresh-cut fruits and vegetables • Salting • Separation and concentration • Smoking • Solar drying • Spray drying • Sterilization (aseptic, retort, home canning) • Syruping • Thermal processing • Use of plastic crates, pre-cooling and cold storage in fresh produce • Vacuum frying • Vacuum packaging	• commercial* channel • confectionery • food loss • food market • food processing technology • food retail • food safety • food waste • grocery store • handling • manufacture • market linkage • market modern* • minimal processing • packaging material • phytosanitary • post-harvest processing • post-harvest storage • processing • processor • raw material • ready eat meal • retailing • shelf stable • supplier • supply chain • transportation • warehousing
OTHER	• Controlled environment agriculture • Vertical farming/vertical agriculture • Hydroponics • Aquaponics • Aeroponics • Nanotechnology • Nuclear techniques/technology • Isotope/Isotopic • Radioisotope • Radiation technique/technology • Food irradiation • Irradiation • Nuclear tracer techniques	

Annex B
Methods

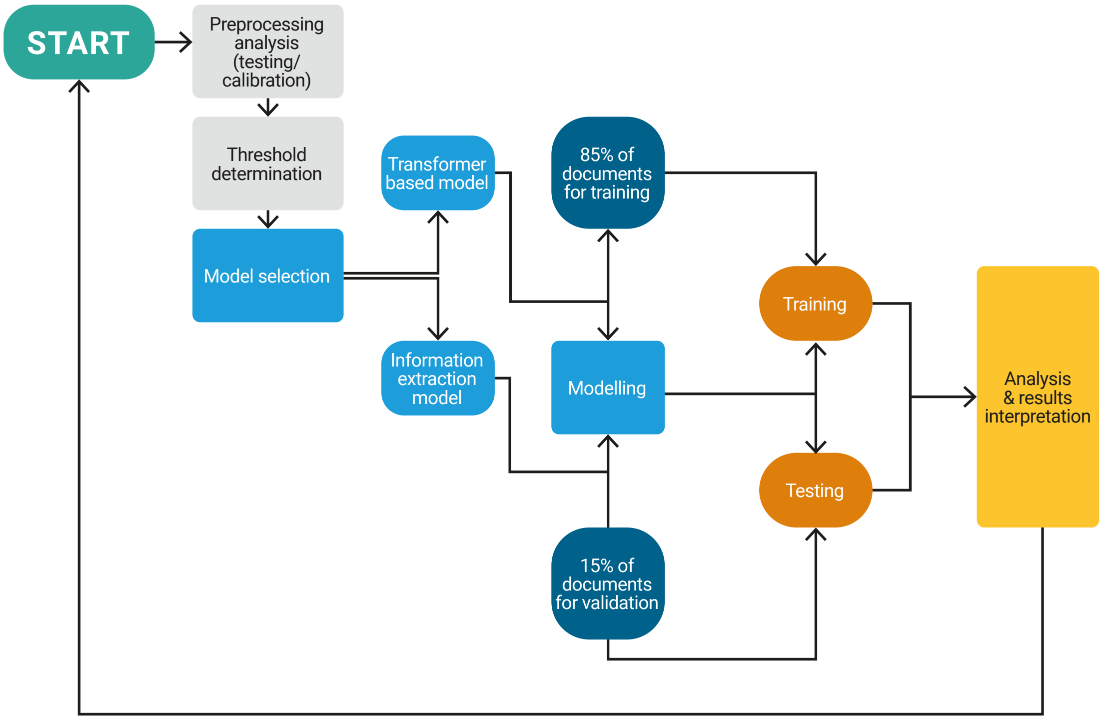
Using training data assembled from collaborative coding from previous exercises and FAO experts, ML was used to support labelling documents, technology, intervention, outcomes and user groups (Figure 1). Figure 14 shows the general workflow to process and categorize documents.

A transformer ML model-based approach was applied to support identification and labelling of document and technology types, outcomes and communities. BERT is a deep-learning language model trained using semi-supervised learning to work with large-scale, unlabelled text by jointly conditioning on both left and right context in all layers (Devlin, 2018). The model evaluates concepts going in both directions, enabling more dynamic and nimble analysis that can be fine-tuned on downstream tasks, achieving state-of-the-art results across a range of natural-language processing tasks. It is designed for sentence-level and token-level tasks and can be applied to the unstructured text.

Several BERT models (base BERT, Roberta, Albert, SciBERT, DistillBERT) were checked for model quality with the settings: 128, 256, 384 and 512 tokens. DistilBERT Named Entity Recognition (NER) uses the BERT architecture but performs knowledge distillation during the pre-training, allowing for lighter, faster and cheaper transformer models, and reduces the size of a BERT model by 40 percent. Due to a limited number of labelled datasets, models were trained by freezing all layers (which is responsible for encoding the text) except the last two layers (where classification occurs). The DistilBERT with two 512 tokens gave the best result with F1-measure = 0.735 on validation ('accuracy' = 0.887, 'F1': 0.735, 'precision': 0.836, 'recall': 0.693). 93 percent of articles' abstracts consisted of fewer than 512 tokens. Thus, the model will not lose very much by restricting training only to the first 512 tokens.

FIGURE 13

Workflow for machine learning



Document and technology labelling

A two-pronged approach was used to support labelling documents with a primary technology, primary document category and multiple outcome labels.

The technology types were used to identify keywords and example technologies associated with that technology type. Additional keywords were produced to supplement expert-derived keyword lists using a customized synonyms database supported by a Word2Vec model (Bojanowski *et al*., 2017). The extended keywords are more general than the expert-produced keywords to ensure a broader suite of possible technologies could be captured. A full list of keywords for each technology is given in Annex A.

The purpose of using keywords as an initial input for each technology type was to ensure that a wide enough net could be cast for analysis because many documents had only title and subtitle data for materials originally produced in English, or materials translated into English. Nearly every entry had a title, whereas summary (abstract) data was provided for about 25 percent of the dataset (Figure 3 in main report). Keywords were useful to support additional training of the ML model.

The training dataset (11 000 documents) was gathered via implicit labelling by searching expert keywords in the documents. Those articles were obtained from the first FAO export, where only 5 percent of articles had an abstract and were used to train and test the model. Eighty percent were used for training, and 20 percent were held aside for testing. A cross-validation technique was used on the training dataset of titles and abstracts. All titles were converted to the vector representation using SentenceTransformer library. An effort was made to use several available pre-trained models, such as 'all-distilroberta-v1", 'all-MiniLM-L12-v2", "all-mpnet-base-v2".

Vector representation was used to support the multi-class classification task. Some classical ML approaches that have low variance, such as RandomForest and LogisticRegression with regularization, were used to ensure that the models were not overfitted. Hyperparameters were chosen based on GridSearch, with the best result obtained by the following combination:'all-MiniLM-L12-v2" + LogisticRegression (with penalty equals to elasticnet and l1_ratio equals to 0.4). A confidence score of 0.75 was identified during document labelling and expert review to offer accuracy thresholds for the model to accept a keyword-based label or a machine-based label. If a machine-modelled threshold of 0.75 was not achieved, then the document was categorized using a keyword function. If no keyword was identified from the expanded synonym list, then a label of "no technology found" was applied.

Technology types

A single label was provided for technology types using the above approach of ML and keyword-based thresholds.

Table 3 provides precision, recall, and F1 measures for technology labelling tasks across the test dataset (1 568 documents) for the best model selected based on cross-validation.

TABLE 3

Precision, recall, and F1 measures for technology labelling tasks for test dataset

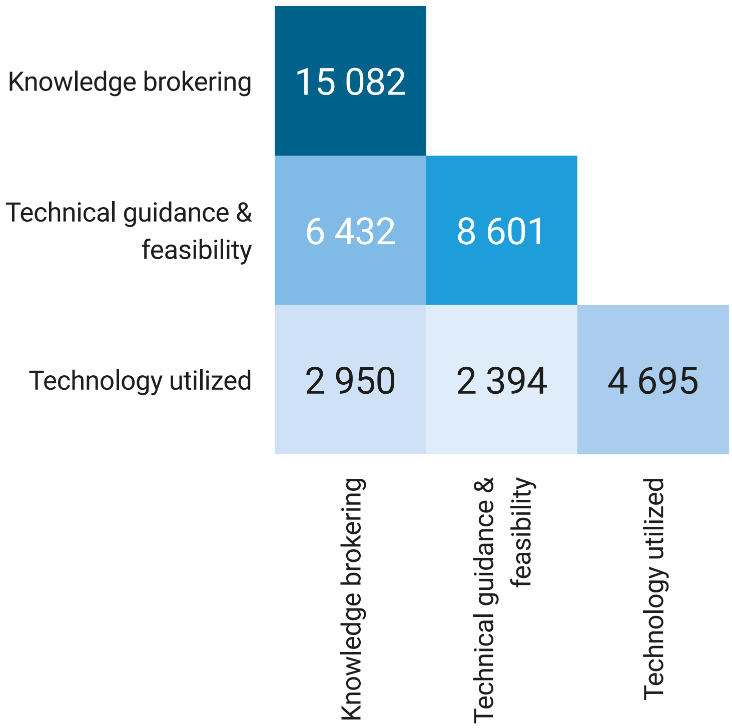
	Precision	Recall	F1-measure	Number of test samples
Biotechnology	0.92	0.90	0.91	429
Digital technology	0.92	0.94	0.93	321
Mechanization*	0.96	0.84	0.89	80
Irrigation	0.90	0.90	0.90	249
Renewable energy	0.88	0.93	0.90	438
Food processing	0.90	0.75	0.82	51
accuracy			0.91	1 568
macro avg.	0.91	0.87	0.89	1 568
weighted avg.	0.91	0.91	0.91	1 568

*mechanization had relatively few test samples available.

Document types

A multi-label approach was used across document categories, where co-occurrence is evident across all categories. A multi-label approach is useful to encourage the model to provide more information about the document and not less, reducing potential bias and siloing resulting from binary classification. A co-occurrence is calculated when there is less than 10 percent difference between the labels. Definitions for document types can be identified in Figure 2 – document category types based on expert elicitation exercises.

FIGURE 14
Numbers of documents for categories of co-occurrence



Outcome trends

The model was trained to extract and label outcomes based on an general outcome typology provided in Tables 4 and 5.

An outcome label was applied to extracted spans of text, and a single document could have multiple outcome spans. This approach was preferred to a single-label outcome with a corresponding threshold because outcome descriptions were typically generalized in the text and only summary text was available. An outcome trend is indicative of intent, and not whether or not an outcome has been achieved.

The co-occurrence highlights where overlap between outcome types occurred.

TABLE 4

Outcome grouping

Parent class	Explanation	Specific outcomes	Definition
Economic growth	Growth across all agriculture or food systems sectors and subsectors that improve the lives of farmers and food systems actors and their families through increases in income, productivity, employment, and practice change	Income amount	Change in income
		Income diversity	Change in sources of income
		Productivity	Change in on-farm crop, labour or livestock productivity or value-chain productivity
		Yield	Change in yield from crop, livestock or foraging
		Adoption	Change in a user's adoption of management or technology related to other agricultural outcomes
		Market efficiency	Change in decision-making based on available, relevant market information
Resilience & risk	Resilience is the ability of individuals, households, communities, cities, institutions, systems and societies to prevent, anticipate, absorb, adapt and transform positively, efficiently and effectively when faced with a wide range of risks, while maintaining an acceptable level of functioning, without compromising long-term prospects for sustainable development, peace and security, human rights and well-being for all. Can include economic, social or climate resilience	Resilience	Change in capacity to prevent, mitigate and recover from shocks and stressors
		Maintenance of operations	Change in capacity to maintain agriculture or food system operation in face of shock
		Community cohesion	Change in vulnerability to conflict, stronger social networks and increased collaboration within a community
Inclusivity, women's empowerment & agency	The process of improving the terms of participation in society, particularly for people who are disadvantaged, through enhancing opportunities, access to resources, voice and respect for rights. This is measured through resulting from the support and inclusive design of all people, but in particular traditionally marginalized groups such as women and people with disabilities, as well as through increased decision-making	Increased knowledge	Change in knowledge about agriculture or food systems related content
		Women's empowerment	Change in women's ability to influence and make decisions independently
		Women's access to resources	Change in women's access to resources (e.g. credit, or inputs)
		Social inclusion	Change in obstacles that limit agency and decision-making capacity
		Gender mainstreaming	Gender mainstreaming into organizational structures and work (NGOs, farmer/village organization)
		Social learning	Change in knowledge and practices through group and community engagement

Parent class	Explanation	Specific outcomes	Definition
Environment & climate action	The process of incentivizing practices that emphasize environmental and planetary health	Environmental sustainability	Change in sustainability of natural resource management such as water, forest or soil management e.g. reduced soil erosion, reduced tree cover loss or increased tree cover,
		Climate mitigation	Change in greenhouse gas emissions
		Change in capacity to adapt to the impacts of climate change	Adaptation and behaviour change that respond specifically to impacts of climate change
		Biodiversity	Change in biological resources at genetic, species or ecosystem level (on-farm or off-farm)
Food security & nutrition	A situation that exists when all people, at all times, have physical, social and economic access to sufficient, safe and nutritious food that meets their dietary needs and food preferences for an active and healthy life. Based on this definition, four food security dimensions can be identified: food availability, economic and physical access to food, food utilization and stability over time. The concept of food security is evolving to recognize the centrality of agency and sustainability.	Dietary diversity	Change in dietary adequacy, including nutrient intake, nutrient adequacy index, and food-based diet quality index
		Food access	Change in an individuals' or households' ability to access food
		Food availability	Change in availability of food
		Malnutrition	Change in malnutrition status
		Nutritious food availability	Change in availability or access to nutritious food
Policy change	Project outcomes explicitly related to change in policies at national, sub national levels	New laws passed or modified	New law passed as per the recommendations of a project, or older laws modified/revised
		Institutional change	Set up of new or modified institutional structures to manage an issue area better

TABLE 5

Outcome precision, recall, and F1 measures

	Precision	Recall	F1-measure
Economic growth	0.83	0.81	0.82
	0.83	0.81	0.82
Inclusivity, women's empowerment & agency	0.72	0.69	0.70
Resilience & risk	0.82	0.78	0.80
Environment & climate	0.79	0.71	0.75
Food security & nutrition	0.84	0.92	0.85
Other outcomes	1.00	0.33	0.50
Policy change	0.81	0.80	0.80

FIGURE 15

Numbers of documents for outcome co-occurrence

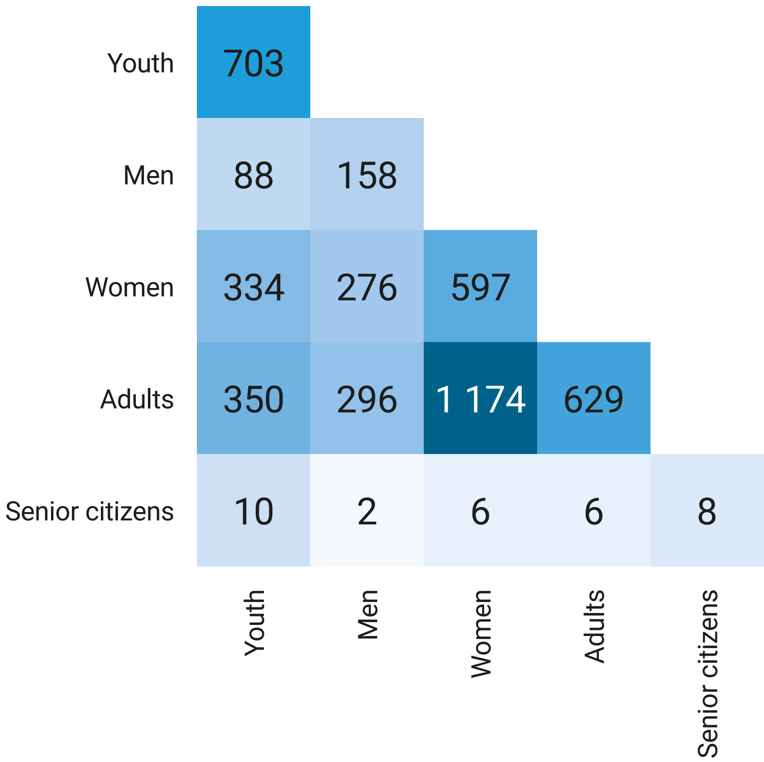
	Inclusivity, women's empowerment & agency	Environment & climate	Economic growth	Resilience & risk	Other outcomes	Policy change	Food security & nutrition
Inclusivity, women's empowerment & agency	482						
Environment & climate	112	1 555					
Economic growth	384	1 272	4 028				
Resilience & risk	44	272	802	808			
Other outcomes	6	26	62	24	59		
Policy change	78	164	554	124	8	682	
Food security & nutrition	156	562	1 864	614	50	252	2 235

Social and demographic details about communities

A pre-trained classifier was used to identify social and demographic details occurring in the text. Information describing populations can be named in various ways and a document may mention details about several populations and/or communities (Figure 16) in the text.

FIGURE 16

Numbers of documents for population co-occurrence



References

Baltenweck, I., Cherney, D., Duncan, A., Eldermire, E., Lwoga, E. T., Labarta, R., Rao, E.J.O., Staal, S. & Teufel, N. 2020. A scoping review of feed interventions and livelihoods of small-scale livestock keepers. *Nature Plants*, 6(10): 1242–1249. **https://doi.org/10.1038/s41477-020-00786-w**

Bojanowski, P., Grave, E., Joulin, A. & Mikolov, T. 2017. Enriching word vectors with subword information. *Transactions of the Association for Computational Linguistics*, 5: 135–146. **https://doi.org/10.1162/tacl_a_00051**

Chandra, R. & Collis, S. 2021. Digital agriculture for small-scale producers: Challenges and opportunities. *Communications of the ACM*, 64(12): 75–84. **https://doi.org/10.1145/3454008**

Devlin, J., Chang, M.-W., Lee, K. & Toutanova, K. 2018. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *ArXiv:1810.04805 [Cs]*. **http://arxiv.org/abs/1810.04805**

Fabregas, R., Kremer, M. & Schilbach, F. 2019. Realizing the potential of digital development: The case of agricultural advice. *Science*, 366(6471). **https://doi.org/10.1126/science.aay3038**

Faccia, A., Cavaliere, L.P.L., Petratos, P. & Mosteanu, N.R. 2022. Unstructured Over Structured, Big Data Analytics and Applications. In: *Accounting and Management. Proceedings of the 2022 6th International Conference on Cloud and Big Data Computing*, 37–41. **https://doi.org/10.1145/3555962.3555969**

FAO. 2021a. Strategic Framework 2022-2031. Rome, FAO. **https://www.fao.org/3/cb7099en/cb7099en.pdf**

FAO. 2021b. FAO Strategy for Private Sector Engagement, 2021-2025. Rome, FAO. **https://www.fao.org/3/cb3352en/cb3352en.pdf**

FAO. 2022a. FAO Science and Innovation Strategy. Rome, FAO. **https://www.fao.org/3/cc2273en/cc2273en.pdf**

FAO. 2022b. Gene editing and agrifood systems. Rome, FAO. **https://doi.org/10.4060/cc3579en**

FAO. 2022c. The State of Food and Agriculture 2022: Leveraging agricultural automation for transforming agrifood systems. Rome, FAO. **https://doi.org/10.4060/cb9479en**

FAO. 2022d. Recent developments in biotechnologies relevant to the characterization, sustainable use and conservation of genetic resources for food and agriculture. FAO Commission on Genetic Resources for Food and Agriculture. Rome, FAO. **https://doi.org/10.4060/cb8956en**

FAO. 2022e. No ordinary irrigation in the Gambia: Innovative solar-powered technologies are securing access to water for rural communities. In: *Food and Agriculture Organization of the United Nations*. Cited 21 February 2023. **https://www.fao.org/fao-stories/article/en/c/1542923/**

FAO. 2022f. Introducing the Agrifood Systems Technologies and Innovations Outlook (ATIO). Rome, FAO. **https://doi.org/10.4060/cc2506en**

FAO. 2023. The status of women in agrifood systems. Rome, FAO. **https://doi.org/10.4060/cc5343en**

Garbaro, A., Carneiro, B., Charpenteir, A., Cerulli, G., Porciello, J., Ivanina, M. & Resce, G. 2020. *Leveraging Artificial Intelligence and Big Data for IFAD 2.0*. **https://www.ifad.org/documents/38714170/42863493/athena.pdf/68906a2a-afde-7f9c-02b7-46822770611b**

Gil, Y., Greaves, M., Hendler, J. & Hirsh, H. 2014. Amplify scientific discovery with artificial intelligence. *Science*, 346(6206): 171–172. **https://doi.org/10.1126/science.1259439**

GSMA. 2022. The State of Mobile Internet Connectivity 2022. London, GSMA. **https://www.gsma.com/r/wp-content/uploads/2022/12/The-State-of-Mobile-Internet-Connectivity-Report-2022.pdf?utm_source=website&utm_medium=download-button&utm_campaign=somic22**

Justice, S., Flores Rojas, M. & Basnyat, M. 2022. *Empowering women farmers - A mechanization catalogue for practitioners*. Rome, FAO. **https://doi.org/10.4060/cb8681en**

Lipper, L., DeFries, R. & Bizikova, L. 2020. Shedding light on the evidence blind spots confounding the multiple objectives of SDG 2. *Nature Plants*, 6(10): 1203–1210. **https://doi.org/10.1038/s41477-020-00792-y**

Liverpool-Tasie, L.S.O., Wineman, A., Young, S., Tambo, J., Vargas, C., Reardon, T., Adjognon, G.S., Porciello, J., Gathoni, N., Bizikova, L., Galiè, A. & Celestin, A. 2020. A scoping review of market links between value chain actors and small-scale producers in developing regions. *Nature Sustainability*, 3(10), Article 10. **https://doi.org/10.1038/s41893-020-00621-2**

McCallum, A. 2005. Information Extraction: Distilling structured data from unstructured text. *Queue*, 3(9): 48–57. **https://doi.org/10.1145/1105664.1105679**

Munn, Z., Peters, M. D. J., Stern, C., Tufanaru, C., McArthur, A. & Aromataris, E. 2018. Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. *BMC Medical Research Methodology*, 18(1): 143. **https://doi.org/10.1186/s12874-018-0611-x**

Nature, 2020. Ending hunger: Science must stop neglecting smallholder farmers. *Nature*, 586(7829), Article 7829. https://doi.org/10.1038/d41586-020-02849-6

Panetta, K. 2021. Gartner Top 10 Data and Analytics Trends for 2021. *Gartner*, March 15, 2021. https://www.gartner.com/smarterwithgartner/gartner-top-10-data-and-analytics-trends-for-2021

Piñeiro, V., Arias, J., Dürr, J., Elverdin, P., Ibáñez, A.M., Kinengyere, A., Opazo, C.M., Owoo, N., Page, J.R., Prager, S.D. & Torero, M. 2020. A scoping review on incentives for adoption of sustainable agricultural practices and their outcomes. *Nature Sustainability*, 3(10): 809–820. https://doi.org/10.1038/s41893-020-00617-y

Pingali, P. 2015. Agricultural policy and nutrition outcomes – getting beyond the preoccupation with staple grains. *Food Security*, 7(3): 583–591. https://doi.org/10.1007/s12571-015-0461-x

Porciello, J. & Ivanina, M. 2021. *The Role of Machine Learning in Programmatic Assessment: A Case Study from USAID's Center for Water Security, Sanitation, and Hygiene*. Research Technical Assistance Center. https://www.rtachesn.org/resources/machine-learning-for-programmatic-synthesis

Porciello, J., Coggins, S., Mabaya, E. & Otunba-Payne, G. 2022. Digital agriculture services in low- and middle-income countries: A systematic scoping review. *Global Food Security*, 34: 100640. https://doi.org/10.1016/j.gfs.2022.100640

Reardon, T., Echeverria, R., Berdegué, J., Minten, B., Liverpool-Tasie, S., Tschirley, D. & Zilberman, D. 2019. Rapid transformation of food systems in developing regions: Highlighting the role of agricultural research & innovations. *Agricultural Systems*, 172: 47–59. https://doi.org/10.1016/j.agsy.2018.01.022

Ricciardi, V., Wane, A., Sidhu, B. S., Godde, C., Solomon, D., McCullough, E., Diekmann, F., Porciello, J., Jain, M., Randall, N. & Mehrabi, Z. 2020. A scoping review of research funding for small-scale farmers in water scarce regions. *Nature Sustainability*, 3(10), Article 10. https://doi.org/10.1038/s41893-020-00623-0

Stathers, T., Holcroft, D., Kitinoja, L., Mvumi, B. M., English, A., Omotilewa, O., Kocher, M., Ault, J. & Torero, M. 2020. A scoping review of interventions for crop postharvest loss reduction in sub-Saharan Africa and South Asia. *Nature Sustainability*, *3*(10): 821–835. https://doi.org/10.1038/s41893-020-00622-1

Takeshima, H. & Vos, R. 2022. *Agricultural mechanisation and child labour in developing countries. Background study.* Rome, FAO. https://doi.org/10.4060/cb8550en

United Nations. 2019. Independent group of scientists appointed by the UN Secretary-General, Global Sustainable Development Report (2019). The Future is Now – Science for Achieving Sustainable Development. https://sdgs.un.org/sites/default/files/2020-07/24797GSDR_report_2019.pdf

Glossary

Key terms and definitions

Key term	Definition
Natural language processing (NLP)	Natural language processing (NLP) is a branch of artificial intelligence (AI) that concerns the interpretation and manipulation of human language by computers. NLP can generate essential-but-currently-missing indicators by performing information extraction and summarization using data from various sources.
Artificial intelligence (AI)	Computer systems that use algorithms to analyse their environment and take actions – with some degree of autonomy – to achieve specific goals. AI can be purely software-based, acting in the virtual world (e.g. voice assistants, image analysis software, search engines, speech and facial recognition systems), or it can be embedded in hardware devices (e.g. advanced robots, autonomous cars, drones or Internet of Things [IoT] applications).
Machine learning (ML)	Machine learning (ML) is a type of AI that uses computer algorithms to automate analytical model building. It is based on identifying patterns in data to improve machine performance by more accurately predicting outcomes without explicit human instructions.
Technology	Technology involves the application of science and knowledge to develop techniques to deliver a new product and/or service or to use a new process to deliver an established product or service. Technologies sometimes emerge serendipitously but are more commonly purposefully developed and are therefore embedded in, and have influence on, social, economic and environmental relations.
Digital technologies	Digital technologies can deliver significant positive impacts, including increased agricultural production and productivity, helping adapt to and mitigate the effects of climate change, supporting early warning systems for plant and animal pests and diseases, improving animal welfare, bringing about more efficient use of natural resources, reducing risk and improving resilience of rural communities, integrating small-scale producers into markets and reaching consumers through e-commerce and increasing efficiency in the design and delivery of agricultural and environmental policies (FAO, 2023).
Biotechnologies	Based on the definition of "biotechnology" in Article 2 of the Convention on Biological Diversity, the term "agricultural biotechnologies" encompasses a suite of technologies from low-tech ones such as artificial insemination, fermentation techniques, biofertilizers and nuclear techniques, to high-tech ones involving advanced DNA-based methodologies (including genetic modification, i.e. GM, genomic selection, whole genome sequencing and gene editing) and multi-omics technologies. They have wide-ranging uses and possibilities including, *inter alia*, crops adapted to biotic and abiotic stresses, nutritionally enhanced and longer lasting foods with reduced losses, reduction of allergens, foodborne disease detection, food safety surveillance, monitoring of genetic diversity and biodiversity, phytoremediation and improved soil health, efficient use of nutrients in feed by animals, rapid diagnosis of diseases and development of vaccines (FAO, 2023).

Mechanization	Mechanization includes farming and processing technologies, ranging from basic hand tools to more sophisticated and motorized equipment. Sustainable agricultural mechanization can reduce drudgery, relieve labour shortages, create new jobs, improve productivity, reduce harvest costs, improve resource use efficiency and enhance market access (FAO, 2023).
Irrigation technologies	Irrigation technologies encompass techniques, skills, methods and processes used to apply water artificially to assist in the growing of crops and pastures. This can be done by letting water flow over the land (surface irrigation), by spraying water under pressure (sprinkler irrigation), or by bringing it directly to the plant (localized irrigation) (FAO, 2023).
Renewable energy technologies	Renewable energy technologies use wind, ocean, solar, hydrological, geothermal and bioenergy sources to generate energy. Transitioning towards energy-smart agrifood systems that optimize the use of efficient and sustainable energy is crucial. Energy-smart agrifood systems not only conserve energy but can even produce it to leverage the dual relationship between energy and food (FAO, 2023).
Food processing technologies	Food processing technologies use methods and techniques involving equipment, energy, and tools to transform agricultural products such as grains, meats, vegetables, fruits and milk into food ingredients or processed food products (FAO, 2023).

Source: FAO. 2023. Technology for transformation of agrifood systems. In: FAO Chief Scientist Office. Rome. Cited 21 February 2023. https://www.fao.org/science-technology-and-innovation/technology/en

ISBN 978-92-5-137967-7

9 789251 379677

CC6724EN/1/08.23