

December 2023

منظمة
الأغذية والزراعة
للأمم المتحدة联合国
粮食及
农业组织Food and Agriculture
Organization of the
United NationsOrganisation des
Nations Unies pour
l'alimentation et
l'agricultureПродовольственная и
сельскохозяйственная
организация
Объединенных НацийOrganización de las
Naciones Unidas para la
Agricultura y la
Alimentación

AFRICAN COMMISSION ON AGRICULTURAL STATISTICS

Twenty-eighth Session

Johannesburg, South Africa: 4 – 8 December 2023

AGENDA ITEM 9

Modernization of dissemination programs with application to agricultural survey programs

Marie Vander Donckt

FAO

INTRODUCTION

Dissemination of official statistics is one key mandate of the agencies of National Statistical Systems (NSS), as pointed out in the United Nations Fundamental Principles of Official Statistics (UNGA 2014). These principles highlight that data produced within the NSS should be considered as a public good and be made accessible simultaneously to the users to enable research, policy analysis and monitoring, and evidence-based decision-making.

With the advent of information and communication technologies, the operational landscape of official statistics has been rapidly evolving over the last two decades. When it comes to dissemination, the new technologies have opened increased possibilities for providers and users of statistic, which at global level has led to an explosion not only in the volume of data available but also in the demand for it.

On the user-side, along with the growing capacity to process large databases using software's, the demand for data has increased alongside with ever more composite expectations: data must be available at a faster rate and in formats addressing different target groups, data must be customizable, data must be comparable and linkable across domains, data must be available at a high level of detail to enable granular analyses.

On the side of producers, IT-developments offers new solution to statistical organizations to move towards more versatile, interactive, impartial, timely and cost-effective release of data. Web-based technologies enable statistical agencies to provide greater access to statistics and metadata and to do so at a lower cost. The detail made available is no longer limited by the size of the publication as it was for print publications. However, to ensure relevance and to take advantage of the modern technologies, statistical organizations must make sure that their dissemination services and tools also evolve over time. In some cases, this may require significant redesign of their current dissemination strategies.

The deliberate and foresighted revamping of dissemination programs is even more critical in a context of constrained human and technical resources that may jeopardize the feasibility and sustainability of these dissemination programs.

This paper reflects upon modernization of dissemination programs for agricultural surveys while weighting the growing range of expectations against constrained human and technical capacities.

DISSEMINATION PROGRAM COMPONENTS

In the context of 50x2030 Initiative¹, the statistics division of the Food and Agriculture Organization of the United Nations (FAO) supports countries at establishing robust national agricultural data systems. One major area of technical assistance bears on strengthening and modernizing the dissemination strategies and programs of national statistics and line ministries in charge of agricultural survey programs. Specifically, FAO works with partner countries at opening and improving access to agricultural survey data and statistics by:

- Promoting data dissemination policies and programmes as part of institutional processes,
- Improving dissemination practices for greater findability and usability of the data,
- Increasing accessibility to agricultural statistics through systematic dissemination, including survey microdata files.

Building on initial assessments of national data dissemination policies and programs for agriculture statistics, FAO provides customized support to improve the existing dissemination protocols and practices, helping to ensure dissemination programs become part of national institutional processes and are organized in accordance with international standards and best practices in the NSO community. For each annual survey cycle, FAO works with the national agriculture statistics agency towards the dissemination of a set of release outputs that balances the needs of the users in a context of limited resources while make best use of web-based dissemination channels. The list of proposed outputs builds around:

1. A statistical bulletin with main results from the annual agricultural survey available.
2. Interactive statistical tables released on a data portal, in compliance with Open Data standards.
3. Dissemination of agricultural survey microdata and associated metadata.
4. Data highlights for wide sharing through instant communication channels and the media.

Taking steps to ensure the viability of an ambitious dissemination program centered around several release products is essential, and this may involve overhauling some more traditional outputs such as reports, tabulation plans, etc.

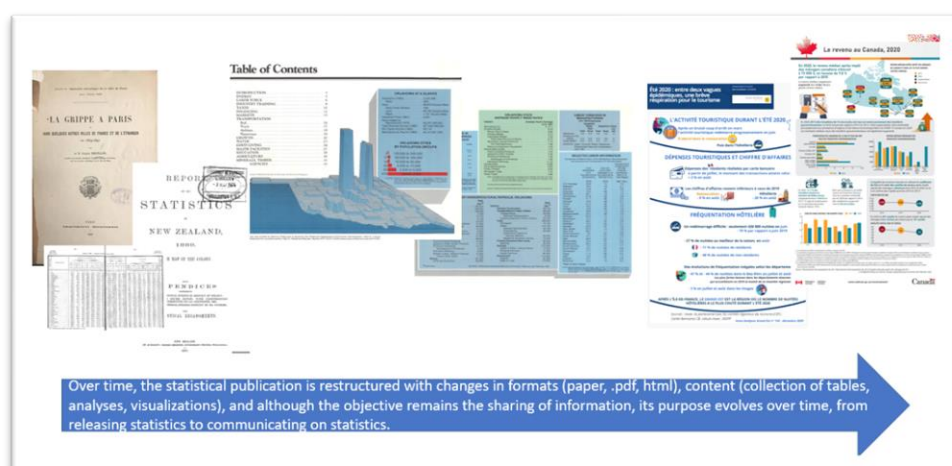
1. Statistical bulletins presenting main results from the agriculture survey.

Increasingly, printed publications are primarily used for display purposes at conferences and events, or as giveaways, but this is a small-scale production of well-designed infographics (e.g., pamphlets). The overall trend is that producing large numbers of printed publications that are essentially a series of tables is declining and being replaced by online access to data, preferably machine-readable files easy to download also in standard formats such as CSV or XLSX.

The recommendation is to move from a “classic” survey report of sometimes hundreds of pages to one or a set of “nutshell” documents that can inform and support debates on specific issues. This will alleviate the burden and free staff time that can be redirected to complementary release outputs such as statistical tables for release on interactive data portals for instance.

¹ The 50x2030 Initiative seeks to empower and support 50 low- and lower-middle-income countries by 2030 in establishing robust national agricultural data systems. These systems will generate high-quality and timely agricultural survey data to inform policies, enhancing countries capacity to produce, analyze, and utilize data for decision-making in the agricultural sector. The FAO oversees the implementation of the data production component, providing technical assistance to countries to support with the adoption of national integrated agricultural survey programs. Source: <https://www.50x2030.org/>.

Figure 1. The evolution of written-type publications by National Statistics Offices.



2. Interactive statistical tables released on a data portal, in compliance with Open Data standards.

In the digital age, data portals have become critical in organizing the sharing of statistical data and metadata. A data portal is a web-based, interactive data and metadata platform with databases modelled for specific data types and domains². The IT-features data portals increasingly support offer many advantages such as:

- Easy access to official statistics and data, downloadable in various data formats, for enhanced usability and decision-making based on data.
- Easy analysis using built-in presentation-ready dashboard and graphics as well as interactive visualization and data extraction tools.
- Compare data across domains through improved harmonization and integration.
- Enhance system interoperability and data exchange among organizations by leveraging on standards such as SDMX to establish a coherent framework for transmitting data and metadata. As data portals are used as data submission tools, the reporting burden by the statistical agencies is reduced.
- Increasingly often, platforms encompass a data management component so that the IT-system is not limited to handling the dissemination of data but enables the management of the wider data production cycle. An example is “*Stat Suite*”, a standard-based, componentized, open-source platform for the efficient production and dissemination of high-quality statistical data. The open-source product is based on the General Statistical Business Process Model (GSBPM) and the Statistical Data and Metadata eXchange (SDMX) standards, and it is the product of a joint effort to co-innovate and co-invest by the open-source community of official statistical organizations in the next generation of practices and solutions to meet shared challenges.³

As technologies evolve rapidly, it is important to keep the platform system management slender. One solution is to avoid maintaining multiple, non-integrated data portals focused on specific domains as may have been done in the past. To the contrary, opting for one common integrated data portal that can be used across an entire national statistical system is a right call. Also, from the users’ perspective, having a one-shop to access the data avoids confusing the users with several portals, some of which may go unnoticed. That is, the integrated data portal increases data usage and does so at a lower cost. Even in decentralized National Statistics System (NSS), it is recommended that all data-producing agencies from the NSS coordinate with the National Statistics Office for a centralized data dissemination where data feed in from the various ministries and other agencies in a standard format such as SDMX.

The African Information Highway (AIH)⁴ - developed by the Statistics Department of the African Development Bank (AfDB), is an example of such an NSS-wide system of data portals that uses SDMX as the common exchange and dissemination format. The AIH is a network of live open data platforms (ODPs) electronically linking all African countries and 16 regional organizations for increased and intelligible public access to official and other statistics across Africa.

² Source: United Nations (2022), “*Handbook on Management and Organization of National Statistical Systems*”(4th Edition, Version 2022/A).

³ Source: <https://siscc.org/stat-suite/>.

⁴ Source: <https://www.afdb.org/en/knowledge/statistics>.

Box 1. Benefits of using Open Data Portals for both improving user experience while simplifying internal processes and streamlining statistical tables compilation by data analysts.

User Experience Improvement:

- Open Data Portals enhance the user experience by providing a platform for presenting complex tables or multi-dimensional data.
- Users can access and analyze data in a more flexible way, making it easier to derive insights and draw conclusions.

Simplification of Internal Processes:

- The use of Open Data Portals can lead to the simplification of internal processes.
- The back-end structure, or the data sets prepared for upload onto the platform, follows a simplified tabulation plan (long format or “flat” tables). This structured approach reduces the workload for data analysts during the compilation process.

Reduction of Workload:

- By structuring data according to a simplified tabulation plan at the back-end level, the workload on data analysts is reduced.
- By presenting multi-dimensional tables on interactive data portals, the inclusion of overly complex or large numbers of tables in reports can be avoided. This contributes to making the information more accessible to a wider audience and to evolve from long statistical reports to the creation of simpler and easy-to-understand tables in reports, such as statistical bulletins and fact sheets.

In summary, the use of Open Data Portals not only benefits end-users by providing a more flexible and user-friendly interface but also contributes to internal efficiency by simplifying data processes and reducing the workload for data analysts. This, in turn, leads to clearer and more understandable reporting structures across dissemination tools such as reports, interactive tables, etc.

3. Dissemination of agricultural survey microdata and associated metadata.

Microdata refers to unit-level data containing information on individual persons, households or business entities collected through a survey or interview or administrative systems.

Microdata dissemination offers a foundation for more nuanced and multidimensional analysis, enabling researchers and policy analysts to test new hypotheses and design programs with greater precision. Beyond these fundamental advantages, releasing microdata provides additional benefits that significantly impact the research landscape. By supporting and diversifying research work, microdata surpasses the limitations of tabular data, which often only answers predefined questions. This approach maximizes the use and usefulness of data, ensuring relevance and reliability through feedback from users. Moreover, the increased return on investments and more efficient use of funding are noteworthy advantages. Transparency and replicability serve as scientific safeguards, with the possibility of replicating aggregates enhancing the robustness of results. As such, it contributes to raising the profile and credibility of the data-producing agency. Furthermore, microdata dissemination can fulfill legal requirements and contractual obligations with sponsors, establishing a comprehensive framework that balances analytical depth with accountability and compliance.

Aware of these benefits, FAO technical assistance programs such as the 50x2030 Initiative supports partner agencies disseminating and promoting user access to these datasets.

Microdata sharing, while offering valuable insights, is accompanied by inherent costs and risks that necessitate careful consideration. Maintaining confidentiality and privacy protection is paramount to uphold trust from respondents, and compliance with both statistical and individual privacy laws is a non-negotiable requirement. The exposure to criticism is an additional challenge, as organizations may find themselves having to respond to inquiries or address potential confusion arising from the distinction between non-official estimates and official estimates. Financial costs present a tangible challenge, encompassing expenses related to data preparation, anonymization processes, and dissemination efforts. Moreover, building and sustaining the necessary technical expertise represent ongoing investments essential for the effective management and secure sharing of microdata. Balancing these costs and risks is crucial to harness the benefits of microdata dissemination while safeguarding the integrity, legality, and public perception of the data-sharing process.

This tension between complying with confidentiality provisions, and balancing the costs and risks, while at the same time opening access to microdata call for the establishment of adequate solutions by NSOs and Line Ministries. These will

suppose working on the key elements constitutive of a microdata dissemination program, namely: the access policy, microdata documentation (including the DDI-standard), microdata anonymization, and the National Data Archive (NADA) cataloguing tool.

Box 2. Key constituents of a microdata sharing program.

The **access policy** refers to the conditions under which microdata are being made accessible must be formal and transparent. Accompanying procedures or protocols, conditions and requirements to fulfill to access the data should be clearly defined in the access policy. The data access policy is interlinked to the access modality (Open Access, Public Use File, Licensed files, data enclaves...).

Statistical Disclosure Control – also known as data anonymization – involves treating a dataset with a view to safeguarding the confidentiality of respondents' identities. This is done by applying technical methods to the datasets aimed at reducing the statistical disclosure risks while balancing the need to preserve data utility.

Metadata (or data documentation) is key to help researchers understand what the data are measuring and how they have been created. Good metadata will lower the risk that the users misunderstand and even misuse the data at hand. It will also reduce the amount of user support statistical offices must offer external users of their microdata. Finally, the use of metadata standards, such as the xml-based Data Documentation Initiative specification (DDI) will enhance data discovery and automated sharing tools, such as survey catalogues that help researchers locate datasets of interest.

The **Microdata catalogue** functions as a repository for micro-datasets, offering users convenient access to both data and accompanying documentation. A robust survey catalogue extends beyond basic storage, providing users with tools to efficiently locate the most pertinent data files for their specific needs. It facilitates the evaluation of information by ensuring compatibility with the researcher's criteria, including considerations of the survey's universe, concepts, and definitions. Moreover, the microdata catalogue serves as a portal for data access through an automated extraction and/or delivery system. An example of such a system is the NADA (Networked Data Archive) system—an open-source software tailored to empower researchers. NADA enables users to effortlessly browse, search, compare, apply for access, and download research data, illustrating the pivotal role of microdata catalogues in enabling streamlined and user-friendly access to valuable datasets.

All the above components interlace and, when approached jointly, may offer solutions to statistical offices to ensure confidentiality safeguards are in place as well as to ensure microdata dissemination programs align on best international standards.

Source: Dupriez, Olivier and Ernie Boyko. 2010. "Dissemination of Microdata Files. Formulating Policies and Procedures", International Household Survey Network, IHSN Working Paper No 005.

4. Data highlights for wide sharing through modern communication channels and the media.

In today's context, an effective communication strategy, centered on highlighting key data points, is essential for the widespread sharing of statistical information. For statistical agencies, integrating instant communication channels and media outreach as central components in dissemination programs is imperative. This approach ensures not only the swift and efficient delivery of crucial insights but also amplifies the dissemination's scope and influence. By utilizing contemporary communication channels and engaging with the media, statistical agencies can convey intricate information in easily digestible formats, promoting broader public comprehension and informed decision-making. As data plays an increasingly pivotal role in shaping narratives and policies, a well-orchestrated communication strategy becomes indispensable, fostering transparency, public awareness, and the overall success of statistical dissemination endeavors.

A GSBPM-BASED APPROACH TO IMPLEMENTING DATA DISSEMINATION.

The nesting of data dissemination activities across all phases of the GSBPM is evidenced in Figure 1. Dissemination activities occur in many sub-processes along the statistics production system and cannot be seen as isolated, unrelated processes, irrespective of the NSS institutional setting. As such, dissemination is to be considered the ultimate objective of a statistical system, with all processes formalized in a dissemination strategy for efficient implementation.

Through the 50x2030 Initiative and the GSARS⁵ project, FAO promotes a GSBPM-based approach to implementing data dissemination.

Achieving sustainability in dissemination programs is crucial for statistical agencies, and this requires careful programming and securing of resources. From this perspective, adopting a GSBPM-based approach to dissemination offers significant advantages. This structured methodology helps statistical agencies mitigate constraints and reduce the risks of falling short of the ultimate objective of any National Statistical Office (NSO), which is the effective sharing of public information. The GSBPM-based approach provides a systematic framework for planning, designing, and executing dissemination processes, ensuring that resources are allocated efficiently, and activities are aligned with the overall objectives. By integrating the GSBPM into dissemination strategies, statistical agencies can enhance their ability to navigate challenges, optimize resource utilization, and achieve sustainable and effective public information sharing.

Figure 2. The Dissemination Phase and interdependencies across the GSBPM.



CONCLUSIONS AND RECOMMENDATIONS

Official statistics serve the crucial role of meeting the evolving and diverse information needs of various user categories. As these needs continually expand and diversify, it becomes imperative to cultivate and enhance a dissemination system that effectively highlights the comprehensiveness and variety of statistical information. This entails not only keeping pace with the changing demands of users but also proactively developing mechanisms to communicate the richness and breadth of available statistical data.

The advancements in information technology provide statistical organizations with novel solutions to transition towards more versatile, interactive, impartial, timely, and cost-effective data releases. Web-based technologies empower statistical agencies to offer increased access to both statistics and metadata at a reduced cost. Unlike traditional print publications, web-based platforms eliminate constraints on the amount of detail that can be made available. However, to fully capitalize on the advantages of these modern technologies, statistical organizations need to ensure that their dissemination services and tools evolve in tandem. This evolution may, in some instances, necessitate a substantial redesign of existing dissemination strategies. Adapting to these technological shifts is essential not only for maintaining relevance but also for maximizing the benefits of a more dynamic and accessible approach to data dissemination.

⁵ Global Strategy to Improve Agricultural and Rural Statistics (GSARS). Website: <https://www.fao.org/food-agriculture-statistics/capacity-development/global-strategy/en/>

In technical assistance programs like 50x2030 and GSARS, FOA strongly advocates for partner statistics agencies to enhance access to data from agricultural statistics programs while leveraging modern technologies for optimization. This entails a strategic reevaluation of dissemination practices, including:

Simplified Reports for Communication: Recommending a shift in reports' purpose towards simplified communication and highlighting key insights rather than just disseminating a collection of tables. This adaptation ensures that reports serve as effective tools for communication on statistics.

Data Portal Distribution: Encouraging the distribution of data through data portals to enhance the user experience. These platforms not only improve accessibility but also can streamline internal data management processes within the agency, contributing to overall efficiency.

Dissemination of Micro-datasets: Emphasizing the dissemination of micro-datasets to maximize their value for evidence-based decision-making. This approach enables further research and analysis, enhancing the agency's credibility and transparency in analytical work.

Strategic Communication Strategy: Proposing the development of a communication strategy centered around the release of key outputs for wider public outreach. This strategy aims to ensure broad awareness of results among the public, fostering a better understanding and appreciation of the agency's contributions.

By adopting these suggested adaptations, partner statistics agencies can not only enhance the accessibility of agricultural data but also leverage modern technologies to streamline processes, maximize the value of datasets, and communicate results effectively to a wider audience.

Achieving sustainability in dissemination programs is crucial for statistical agencies, and this requires careful programming and securing of resources. From this perspective, adopting a GSBPM (Generic Statistical Business Process Model)-based approach to dissemination offers significant advantages. This structured methodology helps statistical agencies mitigate constraints and reduce the risks of falling short of the ultimate objective of any National Statistical Office (NSO), which is the effective sharing of public information. The GSBPM-based approach provides a systematic framework for planning, designing, and executing dissemination processes, ensuring that resources are allocated efficiently, and activities are aligned with the overall objectives. By integrating the GSBPM into dissemination strategies, statistical agencies can enhance their ability to navigate challenges, optimize resource utilization, and ultimately achieve sustainable and effective public information sharing.

REFERENCES

Benschop, T. and Welch, M. (n.d.). “*Statistical Disclosure Control for Microdata: A Practice Guide.*” Retrieved on November 29, 2023, from <https://sdcpractice.readthedocs.io/en/latest/>

Dupriez, Olivier and Ernie Boyko. 2010. “Dissemination of Microdata Files. Formulating Policies and Procedures,” International Household Survey Network, IHSN Working Paper No 005.

United Nations. 2022. *Handbook on Management and Organization of National Statistical Systems.* 4th Edition of the Handbook of Statistical Organization. [New York] Available at <https://unstats.un.org/wiki/display/HSO/Handbook+on+Management+and+Organization+of+National+Statistical+Systems>

United Nations General Assembly (UNGA). 2014. *Fundamental Principles of National Official Statistics. Resolution adopted by the General Assembly on 29 January 2014.* Available at: <https://unstats.un.org/unsd/dnss/gp/FP-New-E.pdf>