UNITED NATIONS ECONOMIC COMMISSION FOR EUROPE

CONFERENCE OF EUROPEAN STATISTICIANS

ModernStats World Workshop 2022

27-29 June 2022, Belgrade, Serbia

Theme: Implementation

INEGI´s Metadata Management Strategy

Authors:

Manuel Cuéllar (INEGI, Mexico), Deputy Director General of Information Integration, manuel.cuellar@inegi.org.mx; José Luján (INEGI, Mexico), Director of Information Governance and Architecture, jose.lujan@inegi.org.mx; Marco Gutiérrez (INEGI, Mexico), Deputy Director of Metadata Management, antonio.gutierrez@inegi.org.mx.

Abstract

It is a fact that standardized metadata is a key factor for the correct and comprehensive use of data. So far, production processes of INEGI have been implemented disjointedly, and with different metadata initiatives, although based on international standards. In this context, and as part of the adoption of the GSBPM, we propose a metadata management strategy to be implemented as a cross-cutting process supported by the coexistence of different metadata standards throughout the entire data lifecycle. Our strategy follows a stages approach: stage 1: definition of a reference metadata specification that allows integral and homologated documentation regardless of data collection method; stage 2: updating the metadata directive; stage 3: designing templates to reuse the standardized evidence of each production process, and stage 4: developing the Metadata Integrator System (MIS) and a metadata management toolkit to facilitate the standardization and systematization of the metadata lifecycle, ensuring interoperability between the different systems and services for metadata consumption.

How was the metadata strategy designed as part of the data governance framework?

Internationally, having standardized metadata has been recognized as a key factor for the correct and comprehensive use of the information produced by the National Statistical Offices. In this context, INEGI's data governance framework includes elements to make it easier for users to access data; standardize data concepts and definitions; promote and facilitate data literacy and improve the quality of statistical and geographical information.

The first step to develop the metadata management strategy was to create or adapt a metadata model that fits the purpose of our organization. Several existing standards and frameworks are available, but there is no one-size-fits-all approach. Not all metadata is the same, for this reason, an important part of our metadata management strategy was to design and categorize metadata specifications by type and purpose.

Another important task was to create and consolidate an area in charge of metadata management. This area is responsible for creating documentation; sharing, disseminating, and archiving metadata policies, and ensuring that the initiatives will be carried out correctly.

> *When a metadata management strategy is developed within the context of a successful data governance framework, the results can be transformative for an organization.*[1]

Our Metadata Management Strategy: management throughout the information lifecycle

INEGI's Strategic Framework points out a paradigm shift to INEGI's business: to complement its traditional role as an information producer with a renewed role of information integrator to facilitate the provision of services with a cross-cutting approach.

Fundamentals of governance, architecture, metadata, and technology infrastructure are extended to the entire data lifecycle: production, integration, and dissemination to consolidate a data ecosystem that satisfies statistical and geographical information needs, and complies with security, confidentiality, quality, and interoperability principles (Figure 1).

---

[1] Varshney, S. (2021, July 27). Data Governance & Metadata Management: Better Together. OvalEdge. https://www.ovaledge.com/blog/data-governance-metadata-management-better-together

Figure 1. INEGI's strategic statistical and geographical information framework
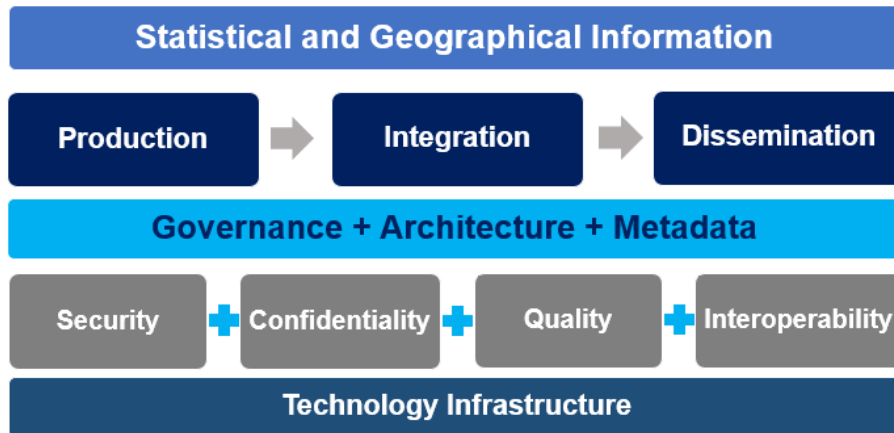
However, due to the diversity of statistical methods and metadata standards, metadata documentation has been implemented disjointedly for each production process, following different metadata initiatives based on international standards.

In this context, and as part of INEGI's Statistical and Geographical Process Model (adapted from GSBPM), we propose a metadata management strategy as a cross-cutting process that allows the coexistence of different metadata standards throughout the entire data lifecycle (Figure 2).

**Figure 2. INEGI's Metadata Management Strategy**
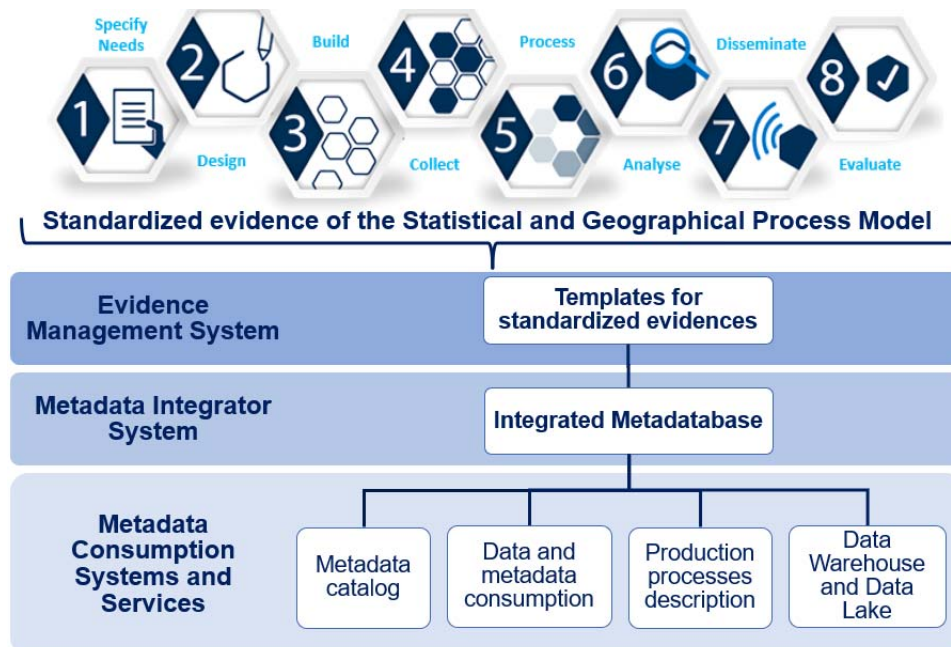
Our metadata management strategy follows a staged approach, starting with:

1. Defining a reference metadata core specification that allows the integral and standardized documentation of the different data collection methods.
2. Updating the technical standard to regulate the documentation activities.
3. Designing templates to reuse documented evidence of each production process.
4. Developing the Metadata Integrator System and a metadata management toolkit that facilitates the capture, updating, sharing, storing, and dissemination.

Metadata Integrator System (MIS)

The Metadata Integrator System (MIS) will consist of four modules. Through a group of services and applications, the MIS will simplify metadata capture, editing, transfer, storage, and dissemination, as well as documentation templates management to standardize and systematize metadata (Figure 3).

Figure 3. Metadata Integrator System
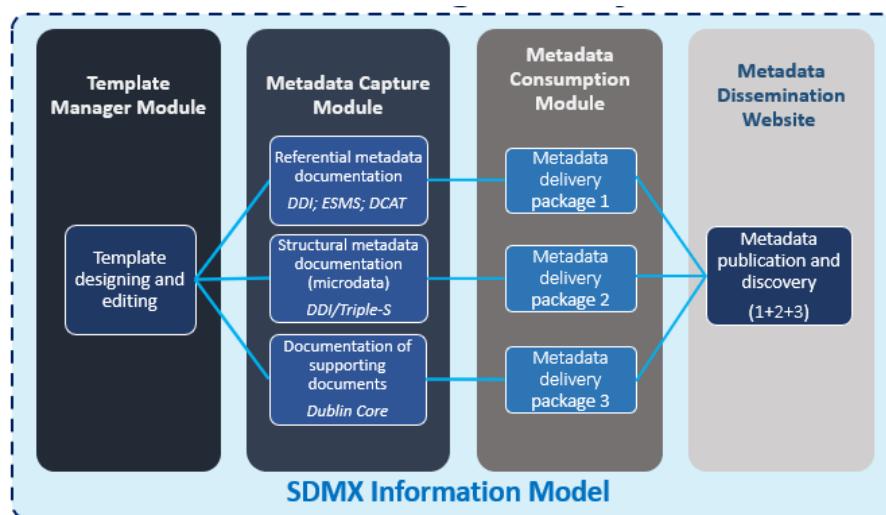
The MIS database design is based on the SDMX information model. This allows the identification of data objects, their relationships, and exchange processes, which will allow centralized and standardized data and metadata management.

MIS' main benefits:

1. Toolkit that satisfies institutional metadata needs and manages metadata workflow.
2. Templates for different metadata specifications according to their purpose and scope.
3. Take advantage of standardized processes' documentation for metadata sourcing.
4. Interoperability between the different services for data and metadata consumption.