

# La publicación de metadatos estadísticos. La experiencia del INE de Uruguay<sup>1</sup>

Lucía Pérez<sup>2</sup>

## Resumen

Los metadatos estadísticos son una herramienta que brinda a los usuarios de los datos, el contexto necesario para su correcta interpretación, detallando información relativa a conceptos, al diseño del proceso estadístico, procedimientos de recolección, procesamiento y análisis de los datos, etc. La existencia de metadatos en una operación estadística mejora la calidad de información estadística que se encuentra disponible y facilita la comparabilidad de datos a nivel internacional.

Para hacer una divulgación adecuada de los datos estadísticos, de forma tal de facilitar el entendimiento de los mismos por todos los usuarios, éstos deben ir acompañados de los metadatos correspondientes. Una vez entendida la importancia de la generación de metadatos estadísticos es que surge la necesidad de estandarizarlos para facilitar el proceso de documentación de los mismos.

Desde 2009 el Instituto Nacional de Estadística (INE) de Uruguay ha trabajado en la documentación de metadatos estadísticos tanto sobre los productos estadísticos propios como de los producidos por las oficinas de estadística del Sistema Estadístico Nacional.

En esta ponencia se expondrá los avances alcanzados por el INE y del Sistema Estadístico Nacional de Uruguay en temas de documentación y publicación de metadatos.

**Palabras clave:** metadatos, estadísticas oficiales, estandarización.

---

<sup>1</sup> Trabajo presentado en las XV Jornadas de Investigación de la Facultad de Ciencias Sociales-UdelaR. Montevideo, 14, 15 y 16 de setiembre de 2016)

<sup>2</sup> Técnico en Gestión de la Calidad – Unidad de Gestión de la Calidad del Instituto Nacional de Estadística del Uruguay. Correo electrónico: [lperez@ine.gub.uy](mailto:lperez@ine.gub.uy)

## Introducción

A partir del último cuarto del siglo XX, se produjo el auge de la informatización y de la socialización de los medios magnéticos, lo que ha facilitado el acceso de una mayor cantidad de usuarios a la información y por consiguiente un aumento en la demanda de datos estadísticos. Es así que se comienza a generar la necesidad de contar con información sobre los datos para hacer una mejor interpretación de los mismos. A esta información se la conoce como metadatos, es decir de una forma sencilla, datos acerca de los datos.

Cuando un usuario hace un análisis de datos de los que no cuenta con información cabal sobre cómo éstos han sido obtenidos puede incurrir en conclusiones erróneas o en un uso inadecuado de los mismos.

Los metadatos estadísticos son una herramienta que brinda a los usuarios de los datos, el contexto necesario para su correcta interpretación, detallando información relativa a conceptos, al diseño del proceso estadístico, procedimientos de recolección, procesamiento y análisis de los datos, etc. Cuando una operación estadística cuenta con metadatos, mejora la calidad de la información que se encuentra disponible y facilita la comparabilidad de datos a diferentes niveles. Es posible clasificarlos en metadatos de referencia, es decir los que contienen información general de la operación estadística y metadatos estructurales que permiten interpretar cada uno de los datos.

Una vez entendida la importancia de la generación de metadatos estadísticos es que surge la necesidad de estandarizarlos para facilitar el proceso de documentación de los mismos.

A nivel internacional existen varios estándares sobre metadatos estadísticos, como por ejemplo el de ISO/IEC 11179 Estándar para el Registro de Metadatos (MDR), es una norma que brinda directrices sobre cómo hacer el registro de datos, su clasificación y su almacenamiento, etc. El estándar SDMX (Statistical Data Metadata eXchange)<sup>3</sup> y el estándar europeo ESMS (Euro-SDMX Metadata Structure) que son iniciativas creadas para facilitar el intercambio de datos y metadatos estadísticos mediante herramientas informáticas. Por su parte en Uruguay se ha optado por seguir el Estándar de la

---

<sup>3</sup> <https://sdmx.org/>

Iniciativa de Documentación de Datos (DDI por su sigla en inglés) para la documentación y publicación de metadatos de las operaciones estadísticas oficiales.

El estándar DDI facilita la documentación y publicación en formato XML de los microdatos y metadatos de las diferentes operaciones estadísticas.

Desde el año 2009 el INE de Uruguay con el apoyo y asistencia del Programa Acelerado de Datos (PAD) ha trabajado en la gestión de los metadatos estadísticos tanto sobre los productos estadísticos propios como de los producidos por las oficinas de estadística del Sistema Estadístico Nacional.

El PAD surge en el año 2006 a partir de las recomendaciones surgidas en el Plan de Acción de Marrakech sobre Estadísticas, un plan de acción global, para lograr un correcto seguimiento de los Objetivos de Desarrollo del Milenio. Este programa busca, junto a la Red Internacional de Encuestas a Hogares (IHSN por su sigla en inglés), aumentar tanto el valor como el uso de los datos estadísticos, haciendo que los datos sean más accesibles, fortaleciendo la demanda y mejorando la calidad de las operaciones estadísticas.<sup>4,5</sup>

El trabajo conjunto de estas dos Organizaciones ha permitido a un gran número de países en África, Asia y el Pacífico, América Latina y el Caribe, como el Uruguay, acceder a herramientas para la estandarización y difusión de datos y metadatos estadísticos fomentando así la mejora de la calidad de los mismos.

Dos de estas herramientas que se han aplicado en el Uruguay han sido el Editor de Metadatos DDI (NESSTAR Publisher 4.0.9) el cual permite documentar los metadatos estadísticos de acuerdo al Estándar DDI y el Catálogo Nacional de Datos (ANDa) para la publicación de microdatos y metadatos.

El editor de metadatos es un programa que ha sido desarrollado por el Centro Noruego para la Investigación de Datos (NDS por su sigla en inglés) de acuerdo a los requerimientos del estándar DDI. El mismo es de uso libre y permite a los productores de estadísticas documentar sus operaciones de una forma sencilla y eficiente. Mediante éste se documentan los metadatos agrupados en 4 conjuntos de metadatos, de acuerdo a la siguiente forma:

---

<sup>4</sup> <http://adp.ihsn.org>

<sup>5</sup> <http://datos.bancomundial.org/>

1. Metadatos que describen el documento de que contiene los metadatos.
2. Metadatos de la operación estadística en sí misma, donde se describen los antecedentes, el diseño muestral (en el caso de encuestas), el procedimiento y herramientas utilizadas para la recolección de los datos, el procesamiento de los mismos, períodos de referencia, indicadores de calidad de la operación estadística, entre otros.
3. Metadatos propios de los microdatos, como ser la identificación de las variables, la definición, tipología y categorías que pueden tomar las mismas, la forma en que son relevadas, la forma en que se relacionan entre sí, si existen datos perdidos, etc.
4. Finalmente los metadatos de los documentos relacionados con la operación que dan contexto a la información a publicar, como pueden ser informes, cuestionarios, manuales de procedimientos de recolección del personal de campo, etc.

Por su parte el programa ANDa, desarrollado por el IHSN, permite publicar tanto los microdatos como los metadatos documentados previamente a través del editor. Mediante este catálogo se logra hacer una difusión eficiente, gracias a su facilidad de uso tanto para los publicadores como de los usuarios de los datos. En él es posible nuclear la información de acuerdo a los temas que abordan, a las instituciones productoras, países, etc., facilitando el acceso de los usuarios a la información.

## **La experiencia en Uruguay**

Desde el comienzo de la cooperación entre el PAD y el INE de Uruguay se ha trabajado haciendo hincapié en la formación de referentes nacionales para fomentar el desarrollo y fortalecimiento de un sistema de gestión de microdatos y metadatos estadísticos a nivel nacional que aumente la accesibilidad y transparencia de los productos estadísticos, mejorando la oferta de los mismo y satisfaciendo la demanda de los usuarios.

En este camino fue que se llevaron a cabo capacitaciones, con instructores internacionales, para concientizar a los referentes temáticos dentro de las distintas oficinas de estadísticas pertenecientes al Sistema Estadístico Nacional (SEN) y que ésta a su vez pudiese ser difundida hacia dentro de la cada oficina.

En esta primera etapa se comenzó por documentar los metadatos de algunas de las operaciones estadísticas más sencillas del INE, se evaluaron las fortalezas y debilidades detectadas durante esta etapa y se establecieron lineamientos para una correcta documentación.

Posteriormente se volvió a realizar otras etapas de capacitación donde además de instructores extranjero se incrementó la participación de nuevos productores y también de usuarios de datos para lograr una mayor difusión de la oferta de datos existentes y de cómo realizar un uso adecuado de los mismos.

Dado que existe un gran número de oficinas productoras de estadísticas pertenecientes al SEN y que se analizó como poco viable el que todas ellas comenzaran la documentación en un mismo momento es que se estableció un conjunto mínimo de metadatos, el cual tenían que cumplir todas las operaciones estadísticas que se publicaran en el Inventario de Operaciones Estadísticas del SEN. A su vez se establecieron otros 3 conjuntos de metadatos más completos, conforme las operaciones estadísticas fuesen: encuestas, censos o registros administrativos. Para la documentación de las operaciones estadísticas del INE se contó con funcionarios propios de la institución, mientras que para la documentación metadatos de otras instituciones, esta fue realizada por consultores internacionales o por sus propios funcionarios previamente capacitados.

Actualmente el INE cuenta en su catálogo de microdatos y metadatos con aproximadamente 20 de sus operaciones estadísticas documentadas en su totalidad y más de 300 operaciones en un nivel primario de documentación correspondiente al inventario antes mencionado.