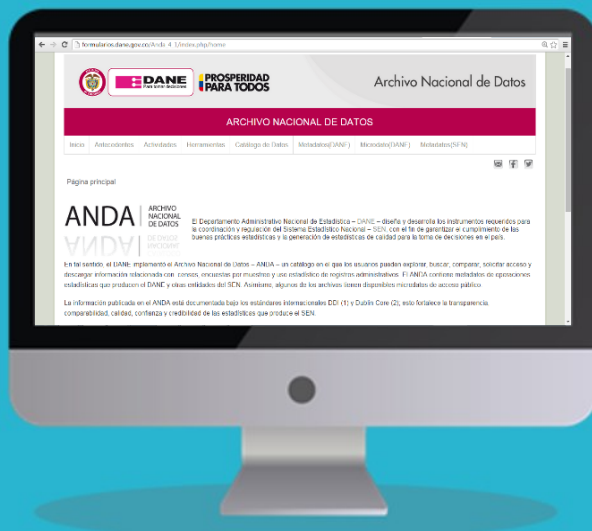


Acceso a información estadística de Colombia en estándares DDI y Dublin Core

Experiencia: Archivo Nacional de Datos - ANDA



http://formularios.dane.gov.co/Anda_4_1/index.php/home



TEMAS CLAVE

Gobierno Abierto (Datos Abiertos)



PALABRAS CLAVE

Estadística, Archivo, Datos, Servicio, Información, Portal Web, Catálogo, Documentación, Microdatos, Metadatos



LOCALIDAD, PAÍS REGIÓN

Colombia, Suramérica



ENTIDADES PARTICIPANTES

Departamento Administrativo Nacional de Estadística: Entidad Nacional, Sector Estadística



ESTADO

Implementación



FECHA DE INICIO IMPLEMENTACION

Noviembre de 2011

El Archivo Nacional de Datos, ANDA, es un instrumento que centraliza los archivos de microdatos y metadatos de las operaciones estadísticas que las entidades producen en Colombia, mediante herramientas sencillas de búsqueda y organización a través de Internet. Este, es un desarrollo que con el apoyo del Banco Mundial, se presenta como pionero en el país y en la región, pues proporciona un repositorio centralizado de datos, que promueve el uso de datos abiertos para la toma de decisiones en política pública, la academia y las oportunidades de negocio. La información almacenada se despliega en diferentes formatos permitiendo la minería de datos.



Justificación, población, objetivo y contexto

LA OPORTUNIDAD QUE SE DESEABA

APROVECHAR:

El Departamento Administrativo Nacional de Estadística (DANE), como ente coordinador y rector de las estadísticas oficiales del país, encontró que algunas de las operaciones estadísticas que realizaban entidades de orden nacional y territorial carecían de las herramientas para normalizarlas, estandarizarlas y ponerlas a disposición del público, junto con sus metadatos.

Por lo tanto, se requería mantener la memoria institucional estadística y poner a disposición del país una herramienta en línea

para acceder a las bases de datos estadísticos.

POBLACIÓN QUE SE BUSCABA IMPACTAR

La población que se quería atender era tanto los productores de información estadística



oficial como los usuarios de dicha información, como por ejemplo los investigadores.

FACTORES IMPULSORES

El DANE promueve la accesibilidad de la información a través de diversos canales, como parte de su política de un DANE moderno; así mismo, está en línea con la implementación de buenas prácticas y con el objeto de producir información que cumpla con los criterios de calidad que se siguen internacionalmente.



Diseño, desarrollo y difusión

EL PROCESO LLEVADO A CABO Y LA SOLUCIÓN

El Departamento Administrativo Nacional de Estadística (DANE) con el apoyo y lineamiento del Banco Mundial, se puso en la tarea de organizar, catalogar y articular un único sistema nacional de estadísticas, donde reposaran los microdatos y metadatos de las operaciones estadísticas generadas en el país.

Como marco de este desarrollo, el DANE adoptó, en el 2008, la iniciativa internacional del Programa Acelerado de Datos – PAD del Banco Mundial. Esta idea surgió en el 2006 como una recomendación del Plan de Acción de Marrakech para las Estadísticas.

EL PAD busca mejorar los programas estadísticos para el adecuado monitoreo de las Metas de Desarrollo del Milenio. Igualmente, tiene como objetivo fortalecer la producción estadística en los países

¹ La Iniciativa de Documentación de Datos (DDI por sus siglas en inglés) es un estándar internacional para describir datos de ciencias sociales, comportamentales y económicas. Expresado en XML, la especificación de los metadatos DDI es compatible con todo el ciclo de vida de los datos de una investigación, incluyendo su

participantes y desarrollar e implementar programas de divulgación y promoción dirigidos a los usuarios de metadatos y microdatos (universidades, centros de investigación, ONGs, etc.).

La finalidad fundamental del PAD es incrementar el uso y valor de la información estadística. En ese sentido, se enfoca en apoyar la implementación de las mejores prácticas y los estándares internacionales sobre documentación y difusión de metadatos y microdatos.

También, hace uso de las herramientas desarrolladas por la Red Internacional de Encuestas de Hogares, que tienen como fin armonizar las prácticas de recopilación y análisis de datos. De este modo, se promueve la comparabilidad internacional de la información y de los indicadores estadísticos.

conceptualización, recolección, procesamiento, difusión, análisis y almacenamiento. Se utiliza para documentar la metodología de la operación estadística, así como las bases de datos.

² *Dublin Core* es un estándar que permite describir e identificar recursos de información. Permite a los usuarios

Con la adopción del PAD, el DANE inició un proceso de instalación del software con los estándares internacionales de la Iniciativa de Documentación de Datos DDI¹ y *Dublin Core*². Se realizaron pruebas al respecto y se adaptaron las plantillas para la documentación.

Para el desarrollo de la solución fue necesario trabajar esencialmente en dos aspectos. Primero, los aspectos asociados con la organización de las estadísticas y segundo, el diseño y desarrollo del software.

Durante el 2009 y el 2010 se recopiló y se documentó la información de las operaciones estadísticas del DANE, comenzando por el año actual y el histórico que permitiera recuperar la memoria institucional. En el 2010, comenzaron a realizarse capacitaciones dirigidas a los productores de información estadística tanto del DANE como de las entidades que conforman el

la búsqueda y recuperación de la información. Se utiliza para documentar material relacionado como manuales, guías, instructivos y demás documentos o reportes generados durante el desarrollo de la operación estadística.

Sistema Estadístico Nacional (SEN)³, en el uso de la herramienta de gestión de microdatos.

Se realizaron capacitaciones en el manejo del software *Nesstar Publisher*, que permite la utilización de los estándares DDI y *Dublin Core* y se instalaron las versiones del software que inicialmente desarrolló el Banco Mundial,

desde que comenzaron las pruebas hasta la presentación de la iniciativa. Con el concurso de diferentes áreas de la entidad, se adaptó el estándar internacional a las operaciones estadísticas que se producen en el DANE, teniendo en cuenta temas como: elementos de formato, creación de catálogos e instructivos, consolidación de información de interés, etc.

Uno de los aspectos más importantes en la implementación de la solución fue el garantizar la anonimización de los datos de las encuestas disponibles, esto con el fin de proteger a las fuentes de información sin sacrificar la calidad y fidelidad del dato estadístico.

En el gráfico 1 se muestra el proceso de producción de un metadato estandarizado, mediante el software *Nesstar Publisher* y los estándares DDI y *Dublin Core*.



Fuente: http://formularios.dane.gov.co/Anda_4_1/index.php/actividades

Una vez documentados los procesos y hechas las respectivas revisiones, la información comenzó a difundirse. En noviembre del 2011 se hizo el lanzamiento del Archivo Nacional de

Datos – ANDA, con el primer set de datos anonimizados, correspondiente a la encuesta de hogares.

Ya implementada la iniciativa se programaron periódicamente capacitaciones en el uso del software para los nuevos usuarios productores de la información que deseaban vincularse.

En definitiva, el ANDA es un catálogo en el que los usuarios pueden explorar, buscar, comparar, solicitar acceso y descargar información relacionada con los censos, encuestas por muestreo y uso estadístico de registros administrativos.

³ El Sistema Estadístico Nacional (SEN) es el conjunto articulado de componentes (productores, instrumentos, fuentes de información, principios, normas, infraestructura humana y técnica, políticas y procesos técnicos) que de

manera organizada y sistemática garantiza la producción y difusión de las estadísticas oficiales que requiere el país. Las entidades que conforman el SEN son los ministerios, departamentos administrativos, entidades

descentralizadas, órganos autónomos y entidades privadas que cumplan con funciones públicas, entre otras.

El ANDA contiene metadatos de operaciones estadísticas que producen el DANE y otras entidades del SEN. Así mismo, tiene disponibles microdatos producidos por el DANE, disponibles para consultas o fines investigativos.

Durante el 2013, se presentó el calendario de divulgación para microdatos y se hicieron actualizaciones de la versión 3.0 a la 3.1. Durante el 2014, comenzaron a difundirse los metadatos de operaciones estadísticas que construyen entidades del SEN. También, se hicieron las actualizaciones de la versión 3.1 a la 4.0. A noviembre del 2014, había un total de 58 operaciones estadísticas del DANE y 6 de entidades externas incluidas en el ANDA.

Entre los beneficiados con esta iniciativa están, por un lado, los productores de información estadística y por el otro, los consumidores de microdatos y metadatos. Los primeros, encuentran en esta herramienta un medio, validado internacionalmente y de gran calidad técnica, para difundir la información que producen.

Los segundos (conformados por investigadores de distintas disciplinas, entidades de gobierno, estudiantes y empresas), pueden documentar de manera rigurosa los procesos con los cuales obtienen información estadística y tienen mayor accesibilidad a los metadatos y microdatos que, además, se encuentran en un lenguaje fácilmente interpretable para emplearlos en investigaciones o análisis de diversa índole.

UTILIZACIÓN DE LA SOLUCIÓN

La difusión de ANDA se hizo en redes sociales, durante el Primer encuentro del Sistema Estadístico Nacional (SEN); en el portal web del DANE; mediante capacitaciones en Bogotá y a nivel territorial del país; y, a partir de sensibilizaciones, a usuarios de metadatos y microdatos.

Los usuarios han mostrado interés y esto se evidencia por medio de los registros de personas que quieren consultar y descargar microdatos. Por ejemplo, en enero del 2013 se encontraban registradas 2.116 personas, a mediados del mismo año la cifra subió a 2.771 y a mediados del 2014 esta cifra se incrementó a 4.969.

La atención a los usuarios que consultan información, permite canalizar las solicitudes y validar la comparabilidad de la información producida.



Impacto, evolución y aprendizajes

PRINCIPALES LOGROS

Se han podido consolidar los metadatos de varias operaciones estadísticas del DANE y de otras varias entidades pertenecientes al Sistema Estadístico Nacional (SEN), y se han ubicado en un único lugar, de fácil acceso para los usuarios, los microdatos de las temáticas social y demográfica del DANE.

EL ANDA es una iniciativa pionera en el país y de gran reconocimiento regional, promueve un gobierno abierto, gracias al uso de los microdatos anonimizados; enseña a las entidades sobre buenas prácticas y cómo debe ser la correcta captura, producción y difusión de estadísticas.

ANDA será una de las variables que permitirá al país su ingreso a la Organización para la Cooperación y el Desarrollo Económicos (OCDE); la información que divulga es de interés para varios sectores del país y los datos que allí reposan no solo permiten la formulación de políticas públicas sino también la creación de negocios, lo cual redundará en bienestar nacional; por último, es una iniciativa democrática pues pueden beneficiarse de ella entidades de orden nacional y territorial.

Los principales beneficios del ANDA se pueden resumir en los siguientes aspectos:

- Contribuye a que los usuarios tengan un mayor conocimiento de la producción de estadísticas en el Sistema Estadístico Nacional, SEN.
- Contribuye a que las estadísticas producidas en el país sean efectivamente aprovechadas en la toma de decisiones.
- Garantiza la difusión de manera estandarizada y organizada de los metadatos de las operaciones estadísticas.
- Garantiza la preservación de los datos y la documentación relacionada además de mantener la memoria institucional.
- Garantiza el acceso equitativo de los usuarios a los metadatos, ofreciendo información detallada de las operaciones estadísticas.
- Fortalece la transparencia, credibilidad y cultura estadística.
- Contribuye a la comparabilidad de las operaciones estadísticas.
- Permite la replicabilidad de los ejercicios estadísticos desarrollados.

LO QUE SIGUE HACIA EL FUTURO

Una de las principales condiciones que permite que esta iniciativa continúe creciendo y fortaleciéndose es el interés que tiene Colombia de ingresar a la Organización para la Cooperación y el Desarrollo Económicos (OCDE), pues iniciativas de este tipo son fundamentalmente para garantizar que el país cuenta con buenas prácticas en materia estadística.

Otra importante motivación que también está asociada a la Estrategia Gobierno en Línea, es que el ANDA promueve y propicia el uso de datos abiertos, con un valor agregado adicional: toda la información dispuesta en esta plataforma cuenta con altísima calidad en los procesos de captura, procesamiento y divulgación de la misma.

Finalmente, el ANDA forma parte de la planeación presupuestal del Departamento Nacional de estadística, lo cual garantiza su sostenibilidad financiera. Además, como parte del propósito superior que se ha trazado el DANE, que es el de contribuir al desarrollo del país produciendo y difundiendo información confiable, relevante, oportuna y de calidad, se busca tener un Sistema Estadístico Nacional cada vez más fortalecido

y que sea derrotero de todas las entidades que producen estadísticas en Colombia, sin importar su sector ni su tamaño.

El principal paso a seguir es tener mayor cantidad de entidades y operaciones en el ANDA.

PRINCIPALES APRENDIZAJES

Los principales obstáculos se encuentran en la falta de información por parte de los productores de estadística, dado que principalmente entidades diferentes al DANE, no tienen la cultura de conservar la memoria

institucional, por lo que se debe adelantar un proceso de recuperación de esta.

Como principal factor de éxito se tiene el crecimiento sostenido en el número de usuarios que consultan el aplicativo y que se registran para descargar microdatos.

COMPARTIENDO LA EXPERIENCIA

Esta iniciativa está siendo implementada por pares a nivel internacional. El caso colombiano está siendo tomado como ejemplo para países de la región quienes están tomando el trabajo de la entidad para

aprovechar la experiencia y así beneficiarse de esta.

Como recomendación en los procesos de implementación de este tipo de iniciativas, se enfatiza en que se tengan en cuenta los ritmos de trabajo de los actores involucrados y que se tenga presente realizar un seguimiento permanente a la documentación para que esta se consolide y se mantenga actualizada.



Información adicional

CONTACTE A LOS PROMOTORES

Erika Mosquera Ortega,

Directora de Difusión, Mercadeo y Cultura Estadística,

Departamento Administrativo Nacional de Estadística (DANE),

(0571) 5978300 Ext 2354,

emosquerao@dane.gov.co.

FUENTES:

- Departamento Administrativo Nacional de Estadística (2014). Postulación Premios Excel GEL 2014, Programa Gobierno en línea, Ministerio de Tecnologías de la Información y las Comunicaciones, Octubre.
- Departamento Administrativo Nacional de Estadística. Archivo Nacional de Datos, ANDA. Visitado el 30 de enero del 2015.
http://formularios.dane.gov.co/Anda_4_1/index.php/home

LOS ARTÍFICES

Estos son los agentes que han hecho parte del proyecto y su contribución.

NOMBRE	ROL
Banco Mundial	Apoyo técnico y financiero para la implementación del ANDA.
Direcciones Técnicas - Departamento Administrativo Nacional de Estadística (DANE)	Dirección del proyecto. Análisis, adaptación de estándares internacionales y software y puesta en marcha de ANDA. Actualización.
Entidades del Sistema Estadístico Nacional (SEN) (Ministerio de Educación, Contaduría General de la Nación, Instituto Nacional Penitenciario y Carcelario INPEC, Instituto Nacional de Salud, Banco de la República)	Publicación de metadatos en ANDA.

LOS HITOS

Los momentos clave de la experiencia presentada son los siguientes:

AÑO	FASES	HITOS
2006	Antecedentes	Nacimiento del Programa Acelerado de Datos – PAD
2008	Diseño	El DANE acoge la iniciativa PAD
		Instalación de Software con los estándares internacionales de documentación de datos (DDI y Dublin Core)
2009 - 2010	Desarrollo	Recopilación y documentación de la información de las operaciones estadísticas del DANE
		Capacitaciones a productores de información (DANE y otras entidades SEN) en herramientas de gestión de microdatos
		Documentación de procesos y revisión información
2011 - 2015	Implementación y mejoramiento	Lanzamiento del Archivo Nacional de Datos – ANDA (Noviembre 2011)
		Capacitaciones periódicas para nuevos usuarios productores de información
		Calendario divulgación para microdatos (2013)
		Actualización versión 3.1 (2013)
		Inicio difusión metadatos entidades SEN (2014)
		Actualización versión 4.0 (2014)