



**Commission on the Status of Women
Sixty-second Session**

**Participation in and access of women to the media, and
information and communications technologies and their
impact on and use as an instrument for the advancement and
empowerment of women**

INTERACTIVE EXPERT PANEL

**Innovative data approaches for measuring progress on
gender equality and women's empowerment**

**Big data y la Modernización de los Sistemas Estadísticos Nacionales:
Experiencia de Colombia**

by

Jaime Sebastian Lobo Tovar *

National Administrative Department of Statistics (DANE),
Colombia

Friday, 16 March 2018

3:00 – 6:00 pm

* The views expressed in this paper are those of the author and do not necessarily represent those of the United Nations.

En Colombia desde 2014 el Departamento Administrativo Nacional de Estadística (DANE) comenzó su proceso de modernización e innovación con el objetivo de impactar en la generación, producción y difusión de estadísticas en Colombia. Esta transformación se desarrolla a través del rol del DANE como coordinador y regulador del Sistema Estadístico Nacional (SEN)¹, sistema que promueve el intercambio de información estadística entre las entidades responsables de registros administrativos a nivel de micro dato como manifiesta la ley 1753 de 2015. En consecuencia, se identifican diversas líneas de trabajo para potenciar la generación y divulgación de información estadística a niveles nacionales y territoriales relacionadas con: (i) la generación de datos inteligentes (Big Data), (ii) capacidad de gestión, (iii) innovación en procesos de apoyo, (iv) datos abiertos y (v) cooperación.

La generación de datos inteligentes hace parte del Plan Estadístico Nacional ² 2017-2022, y es afín con las iniciativas relacionadas con la innovación en procesos de apoyo. Por lo tanto, se reconoce la necesidad de aprovechar fuentes alternativas de información como los datos administrativos; empleando nuevas metodologías que estén la frontera del conocimiento como son los sensores de información geoespacial, la información de telecomunicaciones, la información derivada de transacciones comerciales y bancarias, e información de la web entre otras asociadas al Big Data³. De esta forma, el DANE se adapta a generar estadísticas en un contexto de poca disposición de las fuentes para brindar información, cambios en patrones socio demográficos y una cultura digital que le da un rol cada vez más central a la tecnología.

De este modo, la estrategia de modernización se desarrolla en función de las estadísticas oficiales, con el fin de reducir costos, solucionar problemas presentes en la recolección y procesamiento, desarrollar nuevas metodologías y aprovechar fuentes tradicionales y alternativas de información como el Big Data con el objetivo de fortalecer el Sistema Estadístico Nacional. De esta forma, las fuentes de información asociadas a Big Data pueden facilitar⁴: (i) la comprensión de situaciones en tiempo real, (ii) nueva información sobre movilidad, interacciones sociales, creencias culturales y actividad económica, (iii) alertas tempranas de temas emergentes y crisis, además de (iv) mejoras en la rendición de cuentas y transparencia.

La implantación de *Big Data* se desarrolló como parte del proceso de innovación en los procesos de apoyo, haciendo uso de imágenes satelitales y fotografías aéreas para actualizar el marco censal para el diseño, operación y disseminación de la población de los hogares y la población nacional para el 2018, asimismo se hizo uso de web scraping para mejorar la cobertura y actualizar los datos de identificación y ubicación de las empresas hoteleras. Igualmente, se hizo uso de

¹Sistema Estadístico Nacional: Conjunto articulado de componentes, que de manera organizada y sistemática, garantiza la producción y difusión de las estadísticas oficiales a nivel nacional y territorial que requiere el país. Sus componentes son las entidades u organizaciones que lo integran, usuarios, procesos e instrumentos técnicos para la coordinación, políticas, principios, fuentes de información, infraestructura tecnológica y talento humano.

² Plan Estadístico Nacional: es la **hoja de ruta que tendrá el país para la producción estadística durante los próximos cinco años**. Este documento guiará a las entidades del Sistema Estadístico Nacional en el fortalecimiento de sus actividades estadísticas, indicando estrategias y acciones que permitan, entre otros propósitos, la innovación en la producción y difusión de estadísticas oficiales.

³ Big data es definido por Naciones Unidas como fuentes de datos con un alto volumen, velocidad y variedad de datos, los cuales requieren nuevas herramientas y métodos para su captura y procesamiento de forma eficiente.

⁴ Gender Equality and Big Data. Naciones Unidas. Enero 2018.

scanner data para optimizar el proceso de recolección y ampliar la cobertura de la misma a partir de la recolección centralizada de la información de precios de grandes superficies para el Índice de Precios al Consumidor.

A continuación se describen los proyectos que emplean *Big Data* más relevantes:

Proyecto	Objetivo	Método	Año
Big Data como insumo para la actualización del Directorio Estadístico de Empresas.	Mejorar la cobertura y actualizar los datos de identificación y ubicación de las empresas hoteleras.	Web Scraping	2015-Acualmete
Recolección centralizada de información de precios de grandes superficies para el Índice de Precios al Consumidor- IPC.	Optimizar el proceso de recolección y ampliar la cobertura de la misma a partir de la recolección centralizada de la información de precios de grandes superficies.	Scanner Data	2015-Acualmete
Índice de Precios al Consumidor- IPC	Proponer un método complementario de recolección de información para el IPC, en el que se haga uso de fuentes alternativas.	Web Scraping	2015-Acualmete
Big Data para el cálculo de la relación de la tasa de consumo de suelo urbano y la tasa de crecimiento de población, mediante el uso de imágenes de sensores remotos.	Generar una metodología para el cálculo del Indicador ODS 11.3.1 “Razón entre la tasa de consumo del suelo y la tasa de crecimiento de la población”, a partir del uso de imágenes Satelitales.	Web Scraping	2016-Acualmete
Sistema Integrado de Información sobre Violencias de Género	Disponer, integrar, armonizar y divulgar la información estadística sobre las violencias de género	Datos Administrativos	2016-Acualmete

Incluyendo una perspectiva de género, se reconoce la importancia de la información para mujeres y niñas derivada de fuentes *Big Data* como telefonía celular, sensores remotos y actividad en internet, cuyo contenido puede informar del bienestar socioeconómico y actividad financiera. Por ejemplo, en el caso de la telefonía celular o a través de sensores remotos se puede poseer información acerca del acceso de la mujer al mercado, educación salud y otros servicios esenciales. En el reporte “*Big Data y el bienestar de mujeres y niñas*” reconoce el potencial del manejo de *Big Data* ante las limitaciones de sistemas tradicionales de datos para capturar información detallada de la vida de mujeres y niñas, como un recurso valorable que permite medir el bienestar de mujeres y niñas de forma detallada.

En consecuencia, desde una perspectiva de género a partir de datos administrativos se busca crear el Sistema Integrado de Violencias de Género (SIVIGE) con el objetivo de “*disponer, integrar,*

armonizar y divulgar la información estadística sobre las violencias de género, partiendo de los estándares de calidad, los principios de las estadísticas oficiales y los estándares internacionales; para apoyar el diseño, implementación y evaluación de las políticas públicas y el direccionamiento de las acciones que permitan el abordaje integral de las violencias de género para garantizar el goce efectivo de los derechos”⁵.

El SIVIGE es resultado de un trabajo interinstitucional, en el cual el DANE ha promovido la normalización de conceptos, con el objeto de una lectura global de la *Violencia de Género* por parte de las entidades aportantes de la información estadística. Adicionalmente, el DANE como ente rector, coordinador y regulado del SEN presenta a productores y a usuarios el Código Nacional de Buenas Prácticas para las Estadísticas Oficiales. Lo anterior, para que el SIVIGE como sistema integrado basado en registros administrativos, cumpla con estos principios como sistema y sus fuentes, aquellas que aún no lo han hecho, deban iniciar un proceso de diagnóstico y evaluación para la certificación de sus registros administrativos como estadísticas oficiales.

Por lo tanto, el SIVIGE se convierte en una herramienta para prevenir las violencias y facilitar los procesos que garanticen el goce efectivo de los derechos de las víctimas de violencia de género, siendo una apuesta novedosa para el país, para el continente y para el mundo. Tradicionalmente los institutos nacionales de estadística han optado por la realización de encuestas especializadas, las cuales generan grandes costos; o módulos en encuestas no especializadas que limitan tanto la profundidad de las preguntas pertinentes como el análisis de las mismas. Así, El Sistema Integrado de Información de Violencias de Género ha sido reconocido con mención de honor en la premiación “Interamericana a la Innovación para la Gestión Pública Efectiva” por parte de la Organización de Estados Americanos (OEA), reconociendo el esfuerzo por interoperar la información de violencias de género que recopilan cerca de 21 entidades a nivel local y nacional.

Referencias

Big Data and the Well Being of Women and Girls. (2017). United Nations. Big Data and the Well-Being of Women and Girls Applications on the Social Scientific Frontier

Big Data and Gender. Data2x.org. Econtrado en <<<https://goo.gl/yoMd8c>

Gender Equality and Big Data. (2018). United Nations. Gender Equality and Big Data MAKING GENDER DATA VISIBLE

SIVIGE. (2017) .DANE. Naciones Unidas. UNFPA. Sistema Integrado de Información sobre Violencias de Género

⁵ Marco Normativo, Conceptual y Operativo. ONU mujeres.2017