Procesamiento de datos para la operación estadística: Producción científica bibliométrica en Scopus, Web of Science Core Collection y Scielo Citation Index

Introducción

Este procedimiento busca alinear las actividades con los estándares de calidad y eficiencia, permitiendo optimizar el procesamiento de datos de las bases de datos en Scopus, Web of Science Core Collection y Scielo Citation Index de manera coherente y precisa. Al aplicar estos lineamientos, se garantiza que los cuadros de salida sean los correctos y puedan ser analizados posteriormente. A continuación, se define el objetivo principal que guiará este proceso, el alcance y descripción del proceso.

1. Objetivo

Establecer las actividades necesarias para el procesamiento de datos de la operación estadística de la Producción científica bibliométrica en Scopus, Web of Science Core Collection y Scielo Citation Index.

2. Alcance

Este procedimiento comprende las actividades de la quinta fase: Acopio de los datos, contempladas dentro del modelo del proceso estadístico descritos en los lineamientos para el proceso estadístico en el sistema estadístico nacional, segunda versión y la Norma Técnica de la Calidad del Proceso Estadístico NTC PE 1000:2020. Las actividades contempladas dentro de la quinta fase son: la validación de datos de ingreso, la homologación de etiquetas, la eliminación de registros duplicados base de datos, clasificación y codificación, cálculo de variables por base de datos, integración de las bases de datos, cálculo de variables de base de datos integrada, la revisión y validación, el almacenamiento de la información y por último, la evaluación de la fase de procesamiento.

3. Abreviaturas

OCyT: Observatorio Colombiano de Ciencia y Tecnología.

PEN: Plan Estadístico Nacional

SEN: Sistema Estadístico Nacional.

SGC: Sistema de Gestión de Calidad

4. Definiciones

Para los fines de este documento, se aplican los términos y las definiciones utilizadas en la NTC-ISO 9000, el Decreto 2404 de 2019 y los lineamientos para el proceso estadístico en el Sistema Estadístico Nacional y la Norma Técnica de la Calidad del Proceso Estadístico NTC PE 1000:2020.

5. Lineamientos generales

De acuerdo con lo establecido en el Decreto 2404 de 2019 "por el cual se reglamenta el artículo 155 de la Ley 1955 de 2019 y se modifica el Título 3 de la Parte 2 del Libro 2 del Decreto 1170 de 2015 Único del Sector Administrativo de Información Estadística", el Sistema Estadístico Nacional (SEN) tiene como finalidad establecer e implementar un esquema de coordinación y articulación entre los componentes que lo conforman, que permita mejorar la información estadística producida para la toma de decisiones a nivel nacional y territorial con estándares de calidad, con lenguajes y procedimientos comunes, respetuosos de los estándares estadísticos internacionales y que contribuyan a la transparencia, pertinencia, interoperabilidad, acceso, oportunidad y coherencia de las estadísticas producidas en el país; de modo que la formulación de políticas públicas estén soportadas en evidencia verificables que propenda por una mejora en las condiciones de vida de la sociedad en general.

Por su parte el OCyT en cumplimiento de su objetivo: "fortalecer la capacidad nacional para generar y usar indicadores que servirán para orientar y evaluar las políticas nacionales, regionales e internacionales, así como la acción de diversos actores en los polos científicos y tecnológicos ", ha decidido implementar los requisitos de la NTCPE 1000:2020 con el fin de facilitar que las operaciones del Observatorio hagan parte de las estadísticas oficiales nacionales y territoriales de ciencia y tecnología, así como también, fortalecer la relevancia de la operación para generar información que apoye un mejor seguimiento a las políticas públicas y la toma de decisiones sectoriales.

6. Roles y responsabilidades de los actores participantes en la detección y análisis de necesidades

Rol	Actor	Responsabilidad
Usuario	Son las instancias externas e internas que utilizan la información estadística. Dentro de estos se encuentran ciudadanía en general, entidades públicas, organizaciones privadas, instituciones de educación superior, centros y grupos de investigación, agremiaciones, entre otros, que se consideran relevantes e importantes en el uso de la información estadística para la toma de decisiones.	Manifestar necesidades de información estadística en materia de indicadores bibliométricos.
Entidad aliada	Son las instancias externas que son aliados o socios estratégicos para el desarrollo de proyectos de gestión estadística. Pueden ser entidades públicas o privadas que aportan recursos en el marco de un convenio o contrato.	Cumplir con las obligaciones pactadas mutuamente para el desarrollo del proyecto. Aportar los recursos convenidos y hacer seguimiento al cumplimiento de obligaciones
Socios (Miembros)	Son todas las personas jurídicas que decidan asociarse para el logro del fin común que representa esta persona jurídica, que cumplan con los requisitos, que asuman los compromisos señalados por estos Estatutos y que sean admitidos expresamente por la Junta Directiva.	Asistir a las reuniones de la Asamblea General de Asociados del OBSERVATORIO si les corresponde directamente. Aportar recursos financieros para el desarrollo de la misionalidad del OCyT, de acuerdo con lo defino en los Estatutos. Colaborar con el OCyT en todos aquellos asuntos para los cuales se les requiera, dentro de ellos los relacionados con el desarrollo de proyectos de medición y promover al interior de sus entidades el reporte oportuno de la información requerida para la operación estadística del Observatorio.

Líder de la Política de Calidad Estadística	Director Ejecutivo Subdirector de investigaciones	Dar lineamientos, directrices y políticas para implementar la gestión de la calidad estadística en el OCyT y su mejora continua. Liderar la centralización y difusión oficial de la información estadística que produce y administra el OCyT. Revisar el diseño de las operaciones estadísticas.
Líder funcional del área misional	Líder temático	Verificar y validar el diseño de las operaciones estadísticas. Verificar el procesamiento de la información. Recopilar, validar y consolidar la información recolectada a través de los sistemas de información asociados a las operaciones estadísticas del OCyT. Desarrollar el monitoreo y seguimiento a los indicadores y estadísticas generadas por el OCyT. Analizar la información, interpretar y caracterizar la CteI, para aportar en la toma de decisiones de políticas, planes, programas, estrategias y acciones en la materia. Identificar y gestionar los riesgos de la operación estadística.
Aprobador	Director Ejecutivo o Comité Directivo	Tomar decisiones estratégicas, de lineamientos para la gestión de la información estadística del OCyT. Aprobar la información estadística, antes de su difusión. Liderar el relacionamiento del OCyT con los productores y entidades aliadas de la operación estadística. Velar por la consecución de los recursos requeridos para las operaciones estadísticas del OCyT.

Productor	Son las instituciones, entidades o territorios encargados de la producción de los registros administrativos que se constituyen como el insumo para la producción de la información estadística.	Construir el registro administrativo para producir información estadística con calidad y oportunidad para satisfacer las necesidades de información del usuario.
Líder de Gestión de la Información	Líder de Gestión de la Información	Especificar, diseñar, implementar, disponer y gestionar la infraestructura tecnológica requerida que permita el cargue y procesamiento de los datos producidos, para satisfacer las necesidades de información estadística. Liderar la estrategia tecnológica para facilitar y optimizar la gestión de la información. Liderar la automatización del procedimiento de gestión de la información estadística del OCyT en todas sus fases y la disposición para su difusión Velar por la seguridad informática y la protección de datos acorde con la normatividad y políticas internas vigentes. Implementar nuevas tecnologías para la construcción y publicación de tableros de indicadores amigables, usables y de fácil visualización.
Líder difusión y responsable de atención al usuario	Asistente de comunicaciones	Definir e implementar la política de comunicación interna y externa para una adecuada difusión de la información estadística a los públicos identificados. Gestionar el trámite a las peticiones, quejas y reclamos que se reciban por los diferentes canales definidos por el OCyT. Evaluar la satisfacción de usuario frente a la calidad del servicio prestado por el Observatorio.

Gestor documental y de mejora continua	Líder Calidad	Articular el SGC con los requisitos y criterios de la calidad estadística en el OCyT. Orientar la implementación de los controles a la información documentada generada de las operaciones estadísticas del OCyT. Orientar la implementación de la mejora continua de las operaciones estadísticas. Coordinar el desarrollo de la evaluación de la conformidad de los procesos del OCyT.
Apoyo administrativo	Coordinador Gestión Humana y Administrativa	Brindar el soporte y apoyo administrativo y proveer el recurso humano competente, requerido para la gestión estadística. Liderar las políticas para la gestión documental del OCyT, incluyendo la conservación, custodia y transferencia de los archivos documentales generados de la operación estadística.
Investigador Asistente de Investigación	Investigador Asistente de Investigación	Acompañamiento operativo a la gestión encomendada.
		Apoyo administrativo en la elaboración de ayudas de memoria que surjan en reuniones de trabajo.
		Apoyo a la gestión documental inherente a los procesos estadísticos del OCyT.
		Contribuir en la construcción de la documentación necesaria para que el proceso estadístico del OCyT se presente adecuada y oportunamente a la certificación de la Norma Técnica de la Calidad del Proceso Estadístico (NTC PE 1000) según los lineamientos ofrecidos por el DANE.

		Atender las solicitudes de ajuste y corrección de las entregas documentales que surjan desde la coordinación del proceso de certificación de calidad NTC PE 1000, según los roles y áreas de actuación en las que está involucrado".

7. Desarrollo de actividades

Actividad	Responsable	Punto de control	Descripción	Registro
Validación de datos de ingreso	Ingeniero de soporte Líder temático Investigadores y Asistentes de Investigación	IT-PS-SGC-05 Instrumento de verificación de descargas Archivos procedentes del acopio Aplicación de verificación de etiquetado y formatos IT-PS-SGC-05 Instructivo de procesamiento	Esta actividad verifica la ventana de tiempo, formato, nombres DIVIPOLA	Archivos validados
Homologación de etiquetas	Ingeniero de soporte	Etiquetas de archivos de ingreso IT-PS-SGC-05 Instructivo de procesamiento	Teniendo en cuenta que las etiquetas de las bases de datos tienen diferencia la presentación de los datos, se procede a homologar el etiquetado. Ver instructivo de procesamiento.	Etiquetas homologadas
Eliminación de registros	Ingeniero de soporte	FO-DP-SGI -06 Ficha metodológica	Se realiza una primera depuración y limpieza por base de datos,	Único conjunto de

duplicados por base de datos		FO-DP-SGI-05 Documento metodológico IT-PS-SGC-05 Instructivo de procesamiento	donde se eliminan los documentos duplicados por DOI y luego por título. Los archivos duplicados se eliminan del conjunto de datos y se llevan a la carpeta "Duplicados"	datos por base de datos. Carpeta de archivos duplicados por base de datos
Clasificación y codificación	Ingeniero de soporte	FO-DP-SGI -06 Ficha metodológica FO-DP-SGI-05 Documento metodológico	Se asigna codificación o etiquetado a cada una de las columnas que tendrá los cuadros de salida. Estos códigos y salidas aplican a las bases de datos separadas y a la base de datos integrada.	Definición de códigos en el software e información de cuadros de salida
Cálculo de variables por base de datos	Ingeniero de soporte	FO-DP-SGI -06 Ficha metodológica FO-DP-SGI-05 Documento metodológico IT-PS-SGC-05 Instructivo de procesamiento	Los datos únicos por base de datos son procesados y se generan los cuadros de salida que presentan los indicadores a nivel nacional por base de datos. Esta fase requiere una revisión detallada de resultados antes de presentar los cuadros de salida finales. Se presenta un informe de avance de la fase de procesamiento	Cuadros de salida por base de datos. FO-PS-SGI -02 Informe de avance de la fase de procesamiento.
Integración de las bases de datos	Ingeniero de soporte	Archivos únicos de cada base de datos IT-PS-SGC-05 Instructivo de procesamiento	Esta actividad consolida e integra los datos acopiados en diferentes fuentes. Realiza validación de DOI y títulos, eliminando los registros duplicados, los cuales son enviados a la carpeta "Duplicados"/"Nacional" /" GLOBAL-SCO-WOS-SCI". Se genera un lago único de datos integrados.	Único conjunto de datos integrados Carpeta de archivos duplicados integrados

Cálculo de variables de base de datos integrada	Ingeniero de soporte	FO-DP-SGI -06 Ficha metodológica FO-DP-SGI-05 Documento metodológico IT-PS-SGC-05 Instructivo de procesamiento	Los datos únicos de la base de datos integrada son procesados y se generan los cuadros de salida que presentan los indicadores a nivel nacional.	Cuadros de salida consolidados a nivel nacional .
Revisión y Validación	Ingeniero de soporte Líder temático Investigadores y Asistentes de Investigación	FO-DP-SGI -06 Ficha metodológica FO-DP-SGI-05 Documento metodológico Cuadros de salida de base de datos integrada. FO-PS-SGI -02 Informe de avance de la fase de procesamiento.	Se realiza una revisión detallada de los resultados, analizando las posibles inconsistencias. De existir se realiza el reporte y se ajusta.	Cuadros de salida revisados y validados FO-DP-SGI -03 Informe final de la fase de procesamiento
Cierre de los archivos de datos y Almacenamiento de la información	Ingeniero de soporte Líder temático Investigadores y Asistentes de Investigación	Cuadros de salida revisados y validados	Una vez se tiene finalizada la información se revisa junto al equipo de bibliometría y de estar en completo orden se da por cerrados los archivos de datos y se procede a almacenar en el Drive del OCyT y a realizar copias de seguridad	Información almacenada Acta de cierre de la fase
Evaluación del procesamiento	Ingeniero de soporte Líder temático	Cuadros de salida de base de datos integrada.	Se evalúa el desempeño de los procesos, herramientas, técnicas y métodos usados en la fase.	FO-DP-SGI -03 Informe final de la fase de procesamiento actualizado.

	Investigadores y Asistentes de Investigación	FO-DP-SGI -03 Informe final de la fase de procesamiento	El reporte se agrega al informe final de la fase de procesamiento.	

Listado de documentos

- FO-DP-SGI-05 Documento metodológico
- F-DP-SGI-03 Plan General
- Acta de Reunión
- FO-DP-SGI -06 Ficha metodológica
- IT-PS-SGC-05 Instructivo de procesamiento
- FO-PS-SGI-05 Instrumento de verificación de descargas
- FO-DP-SGI -02 Informe de avance de la fase de procesamiento
- FO-DP-SGI -03 Informe final de la fase de procesamiento
- FO-DP-SGI-03 Informe de evaluación
- Acta de revisión al diseño de la operación estadística

8. Ficha de aprobación

Elaborado por:	Revisado por:	Aprobado por:
Nombre: Alicia Herrera Sánchez **Cargo:** Asistente de Investigación **Fecha**: 11/12/2024	**Nombre:** Efrén Romero Riaño **Cargo:** Líder temática Bibliometría **Fecha**: 11/12/2024	**Nombre:** Andrea Guevara Rey **Cargo:** Líder temática Inversión **Fecha**: 11/12/2024

9. Control de Cambios

FECHA	DESCRIPCIÓN DEL CAMBIO	VERSIÓN
11/12/2024	Lanzamiento de documento	1