DANE
INFORMACIÓN PARA TODOS

SEN 2.0
Sistema Estadístico Nacional-Colombia



Curso Auditor Interno NTC PE1000

Material de Curso Virtual Formación de Auditores Internos NTC PE 1000 (Julio 2021)

Módulo 5
Calidad de los Archivos de Datos
Parte 2

Julio 2021

El futuro es de todos
Gobierno de Colombia

Requisitos inherentes a la calidad de las bases de datos

Requisito 2017:

4.3 CONFIDENCIALIDAD DE LA INFORMACIÓN RECOLECTADA

4.3.1 La entidad debe implementar mecanismos que garanticen que la información recolectada de las fuentes se conserva de manera confidencial conforme a la normatividad vigente de protección de datos personales.

Requisito 2020:

4.3 CONFIDENCIALIDAD DE LA INFORMACIÓN RECOLECTADA

4.3.1 La entidad debe implementar mecanismos que garanticen que los datos recolectados de las fuentes y la información estadística producida, se conserve de manera confidencial conforme a la normatividad vigente de protección de datos personales y los Principios Fundamentales de las Estadísticas Oficiales de Naciones Unidas.



Verificar:

- ✓ Tratamiento de datos personales.
  - ✓ ¿Se han establecido mecanismos para garantizar la protección de información sensible contra divulgación no autorizada, cumpliendo lo establecido en la normatividad de datos personales?

Material de Apoyo Auditores Internos NTC PE 1000 (Julio 2021)

Requisitos inherentes a la calidad de las bases de datos



Requisito 2017:

4.3 CONFIDENCIALIDAD DE LA INFORMACIÓN RECOLECTADA

4.3.2 La entidad debe mantener de forma segura la información recolectada para prevenir el acceso o uso no deseado, daños o pérdida fortuita.

Requisito 2020:

4.3 CONFIDENCIALIDAD DE LA INFORMACIÓN RECOLECTADA

4.3.2 La entidad debe conservar de forma segura los datos recolectados para prevenir el acceso o uso no deseado, daños o pérdida fortuita y la confidencialidad de los datos

Verificar:

- ✓ Si hay una política de seguridad y acceso.
- ✓ Solicitudes de creación de usuarios.
- ✓ Clausulas de confidencialidad en contratos.
- ✓ Uso de ID de usuario y claves de acceso en el software aplicativo.
- ✓ Certificación de la Gestión de Seguridad de la Información ISO/IEC 27001. (no obligatoria)

Material de Curso Virtual Formación de Auditores Internos NTC ISO 9001:2015 (Junio 2021)

Material de Curso Virtual Formación de Auditores Internos NTC PE 1000 (Julio 2021)

Capitulo 5
Detección y analisis de necesidades
NTC PE 1000:2020

Requisitos inherentes a la calidad de las bases de datos

Material de Curso Virtual Formación de Auditores Internos NTC PE 1000 (Julio 2021)

Requisito 2020:

5.6 COMPROBACIÓN DE LA DISPONIBILIDAD DE DATOS 5.6.3 Comprobación de los archivos de datos

Aplicable para operaciones estadísticas que utilicen los registros administrativos o fuentes alternativas de datos para el aprovechamiento estadístico.

Literales: a,g,h, i, j, k

NOTA: los registros administrativos y las fuentes alternativas se encuentran consolidadas en archivos de datos.

Verificar:

- ✓ Verificar como se establece la relación entre las entidades
- ✓ Verificar si cuentan criterios de calidad para evitar la duplicidad de los datos.
- ✓ Verificar si cuentan con un cronograma para el envío de archivos de datos.
- ✓ Verificar como es el esquema de acceso a los datos
- ✓ Verificar el esquema tecnológico para el procesamiento de los archivos de datos
- ✓ Verificación de herramientas para el manejo de confidencialidad, disponibilidad y protección de los datos
- ✓ Indagar si hay acceso continuo a los archivos de datos para la producción estadística

Material de Curso Virtual Formación de Auditores Internos NTC PE 1000 (Julio 2021)

Capitulo 6
Diseño
NTC PE 1000:2020
Diseño y pruebas
NTC PE 1000:2017

Requisitos inherentes a la calidad de las bases de datos



P

Requisito 2020:

6.3.3.1 La operación estadística debe evidenciar el uso de las nomenclaturas y las clasificaciones estadísticas adoptadas, adaptadas o elaboradas en su última versión disponible.

.

Verificar:

- ✓ Verificar si hay variables asociadas a las nomenclaturas y las clasificaciones estadísticas adoptadas, adaptadas o elaboradas en su última versión disponible, indicadas como uso especifico en los documentos metodológicos.

Material de Curso Virtual Formación de Auditores Internos NTC PE 1000 (Julio 2021)

Requisitos inherentes a la calidad de las bases de datos



P

Requisito 2017:

6.5 DISEÑO DE LA EJECUCIÓN

6.5.2 La entidad debe establecer, documentar e implementar las reglas de validación y consistencia.

.

Verificar:

- ✓ ¿La entidad ha establecido e implementado las reglas de validación y consistencia?
- ✓ ¿Las especificaciones y/o reglas de validación se encuentran documentadas para todos los campos de la base de datos?
- ✓ ¿Cuenta con un documento que describa las especificaciones de validación y consistencia?

Material de Curso Virtual Formación de Auditores Internos NTC PE 1000 (Julio 2021)

Requisitos inherentes a la calidad de las bases de datos



Requisito 2020:

6.5.3 Apoyo de datos estructurados

Aplicable para operaciones estadísticas que realizan el acopio de archivos de datos estructurados, como lo son a partir de registros administrativos, derivadas y grandes archivos de datos.

La operación estadística debe determinar y gestionar el intercambio de los archivos de datos que contengan los microdatos y sus metadatos.

Verificar:

- ✓ Documentos asociados a acuerdos a nivel de servicio
- ✓ Documentos técnicos asociados con las condiciones para el intercambio de datos.
- ✓ Verificar que tan formales son estos acuerdos y sus documentos de soporte tales como convenios, acuerdos y compromisos de cada una de las partes.

Material de Curso Virtual Formación de Auditores Internos NTC PE 1000 (Julio 2021)

Requisitos inherentes a la calidad de las bases de datos



Requisito 2020:

6.5.4 Acopio de datos no estructurados

Aplicable para operaciones estadísticas que realizan el acopio de archivos de datos no estructurados, como imágenes, texto, archivos de audio y video, entre otros, que obtienen de diversas fuentes como redes sociales, aplicaciones, sensores remotos, sitios web, etc.

Verificar:

- ✓ Identificar las herramientas tecnológicas y las especificaciones técnicas

NOTA: las herramientas tecnológicas de acopio pueden ser:

- ✓ librerias para web scraping, se implementan para diferentes lenguajes de programación (Python, R, Scala), e
- ✓ interfaces de programación de aplicaciones (API, por su sigla en inglés Application Programming Interface), normalmente las brindan los proveedores de datos para conectarse directamente a las fuentes de datos y extraer la información.

Material de Curso Virtual de Formación de Auditores Internos NTC PE 1000 (Julio 2021)

Requisitos inherentes a la calidad de las bases de datos



Requisito 2020:

6.5 DISEÑO DE LA RECOLECCIÓN O ACOPIO

6.5 La recolección aplica a las operaciones estadísticas con fuentes primarias de datos y el acopio a las operaciones estadísticas con fuentes secundarias.

Verificar:

- ✓ Verificar si la entidad determinó el periodo de tiempo y el medio que se va a utilizar para la conservación de los datos obtenidos.

Material de Curso Virtual Formación de Auditores Internos NTC PE 1000 (Julio 2021)

Requisitos inherentes a la calidad de las bases de datos



Requisito 2020:

6.5.5 La operación estadística debe determinar los criterios para verificar la calidad de los datos acopiados

Verificar:

- ✓ Verificar como la entidad determina los criterios para determinar la calidad de los datos acopiados, como se hace seguimiento y control a esos criterios.
- ✓ Verificar el resultado de las evaluaciones de calidad de datos aplicado a las fuentes específicamente sobre las variables acopiadas
- ✓ Verificar la existencia de reglas de validación de las variables fuente
- ✓ Identificar los métodos que implementan los responsables para verificar la calidad de las variables fuente, de acuerdo a lo identificado y documentado en el requisito 4.4.5.3.

Material de Curso Virtual Formación de Auditores Internos NTC PE 1000 (Julio 2021)

Requisitos inherentes a la calidad de las bases de datos



Material del Curso Virtual Formación de Auditores Internos NTC PE 1000 (Julio 2021)

Requisito 2020:

6.6 DISEÑO DEL PROCESAMIENTO DE LOS DATOS
La operación estadística debe determinar y documentar:

a) la consolidación o la integración de los archivos de datos, incluidos los datos geográficos,
b) la verificación de la completitud de los datos,
c) las especificaciones de rutina para la codificación (cuando sea aplicable),
d) la edición (validación y consistencia de los datos),
g) la validación y consistencia de la información alfanumérica y geográfica (cuando sea aplicable),
i) el medio de acopio o de recepción de los datos recolectados.

Verificar:

- ✓ Verificar la existencia de la documentación técnica asociada a estos literales.

Requisitos inherentes a la calidad de las bases de datos



Requisito 2017:

6.6.1 La entidad debe, de acuerdo con las características de la operación estadística, determinar la infraestructura tecnológica necesaria para ejecutar la operación estadística y definir el aplicativo (software) que se requiere para el desarrollo de las siguientes actividades:

¿La entidad ha determinado la **infraestructura tecnológica** necesaria para ejecutar la OE?

¿La entidad ha definido el **software requerido** para desarrollar las actividades referidas en la NTCPE-1000?

¿Cuenta con los documentos de detección de requerimientos para el desarrollo del software?

¿El **software utilizado es suficiente** para cumplir con los objetivos definidos para la operación estadística?

Verificar:

Verificar:

- ✓ Software aplicativo para la captura y procesamiento de la información.
- ✓ Herramientas de software para el análisis estadístico de la información.
- ✓ Medios de difusión.

Material del Curso Virtual Formación de Auditores Internos NTC PE 1000 (Julio 2021)

Requisitos inherentes a la calidad de las bases de datos



Requisito 2020:

6.6.1 DISEÑO DEL PROCESAMIENTO DE LOS DATOS

La operación estadística debe determinar la infraestructura tecnológica necesaria para la ejecución del proceso estadístico y definir el software aplicativo requerido para el desarrollo de las siguientes actividades:
- la captura, la digitación, la digitalización, la transmisión y la integración de los datos (cuando sea aplicable),
- el acopio de datos,
- el almacenamiento, la lectura y la seguridad de los archivos de datos,
- la aplicación de los procesos estadísticos de detección de inconsistencias, la validación, la corrección y la depuración de datos, la estimación, y
- la generación de resultados, cuadros de salida, análisis de resultados, entre otros insumos para la interpretación de estos.

Verificar:

- ✓ Verificar en los documentos de diseño del procesamiento si esta descrita la infraestructura tecnológica necesaria y software aplicativo para la OE.
  - ✓ La captura, la digitación, la digitalización, la transmisión e integración de los datos (cuando sea aplicable),
  - ✓ el acopio de datos,
  - ✓ el almacenamiento, la lectura y la seguridad de los archivos de datos.

Material de Curso Virtual Formación de Auditores Internos NTC PE 1000 (Julio 2021)

Requisitos inherentes a la calidad de las bases de datos



Requisito 2017:

6.6.2 La entidad debe documentar, implementar, mantener y actualizar, la documentación que se genera para el uso y el mantenimiento del aplicativo (software) que requiera en el desarrollo de la operación estadística, esta documentación debe incluir:

- modelo entidad-relación o modelo de datos,
- diccionario de datos,
- manual de validación y consistencia,
- manual de captura,
- manual de usuario,
- manual del sistema, y
- mecanismos para garantizar seguridad, mantenimiento y respaldo de la información.

Requisito 2020:

6.6.2 La operación estadística debe establecer la documentación para el uso y mantenimiento del aplicativo (software), esta documentación debe incluir:

a) modelo entidad-relación o modelo de datos,
b) diccionario de datos,
c) manual de validación y consistencia (cuando sea aplicable),
d) manual de captura,
e) manual de usuario,
f) manual del sistema, y

g) mecanismos para garantizar la seguridad, el mantenimiento, conservación y el respaldo de los datos.

Verificar:

- ✓ Realizar la verificación de la existencia de los documentos técnicos indicados en este numeral.

Material de Curso Virtual Formación de Auditores Internos NTC PE 1000 (Julio 2021)

Requisitos inherentes a la calidad de las bases de datos



Material de Curso Virtual Formación de Auditores Internos NTC PE 1000 (Julio 2021)

Requisito 2020:

6.7.3 La entidad debe determinar las herramientas tecnológicas requeridas para el análisis de la información estadística o geoestadística obtenida.

Verificar:

- ✓ Verificar si en los documentos técnicos del diseño del procesamiento de la información están indicadas las herramientas tecnológicas requeridas para el análisis de la información estadística o geoestadística obtenida.
- ✓ Uso de SAS, SPSS, R etc.

Requisitos inherentes a la calidad de las bases de datos



Material de Curso Virtual Formación de Auditores Internos NTC PE 1000 (Julio 2021)

Requisito 2020:

6.10 FINALIZACIÓN DE ARCHIVOS DE DATOS

La entidad debe determinar y documentar los criterios para el cierre de los archivos de datos.

El cierre de los archivos de datos aplica a las fases de:

a) recolección o acopio
b) procesamiento
c) análisis

Verificar:

- ✓ Verificar si los criterios de cierre de los archivos de datos están previstos y documentados.

Requisitos inherentes a la calidad de las bases de datos



Material de Curso Virtual Formación de Auditores Internos NTC PE 1000 (Julio 2021)

Requisito 2020:

6.11.2 La operación estadística debe considerar las pruebas funcionales y no funcionales para los aplicativos.

Verificar:

- ✓ Revisar la existencia de procedimientos o documentos tipo registro para el desarrollo de las pruebas funcionales y no funcionales y la aceptación de estas por parte de los usuarios; así como los requisitos de la puesta en producción.

Material de Curso Virtual Formación de Auditores Internos NTC PE 1000 (Julio 2021)

Capitulo 7
Construcción
NTC PE 1000:2020
Ejecución
NTC PE 1000:2017

Requisitos inherentes a la calidad de las bases de datos



Material de Curso Virtual Formación de Auditores Internos NTC PE 1000 (Julio 2021)

Requisito 2017:

7.2 RECOLECCIÓN DE DATOS
7.2.1 Captura y control del flujo de la información
7.2.1.2 La entidad debe consolidar en una base de datos la información recolectada llevando a cabo:
validaciones para revisar, la consistencia de los datos, y
los controles sobre la posible pérdida de información.

Verificar:

- ✓ Verificar aleatoriamente, como operan las validaciones en el Software Aplicativo.

Requisitos inherentes a la calidad de las bases de datos

Requisito 2020:

7.3 DESARROLLO DEL SOFTWARE APLICATIVO DE LA OPERACIÓN ESTADÍSTICA

La operación estadística debe desarrollar y mantener el software aplicativo de acuerdo con las especificaciones del diseño.

NOTA 1: el software aplicativo puede ser desarrollados para el diseño muestral, la recolección, la transmisión, el procesamiento, el análisis, la difusión y sus submódulos.

NOTA 2: la entidad debería desarrollar software aplicativo o módulos para el entrenamiento, el seguimiento, el control y la evaluación del desempeño, así como desarrollar macros para generar reportes o informes y gráficos para el análisis de resultados, entre otros.

NOTA 3: el software aplicativo de transmisión de archivos de datos debería contener campos de identificación geográfica para facilitar la georreferenciación.

Material de Curso Virtual Formación de Auditores Internos NTC PE 1000 (Julio 2021)

Verificar:

- ✓ Verificar que el software aplicativo cumpla con las especificaciones definidas en el diseño.

Requisitos inherentes a la calidad de las bases de datos



Requisito 2020:

7.3.1 La entidad debe implementar las reglas de validación y consistencia en el desarrollo de los aplicativos.

Verificar:

- ✓ Verificar como operan las validaciones sobre los aplicativos de captura y/o acopio de información.
- ✓ • Traducción y ejecución de las reglas de validación para evaluar la calidad de los archivos de datos y analizar los indicadores para el informe de evaluación de calidad..

Material de Curso Virtual Formación de Auditores Internos NTC PE 1000 (Julio 2021)

Requisitos inherentes a la calidad de las bases de datos



Material de Curso Virtual Formación de Auditores Internos NTC PE 1000 (Julio 2021)

Requisito 2020:

7.3.2 Desarrollos para el acopio de datos no estructurados (aplicativos)

La entidad debe determinar el algoritmo de búsqueda, o el script para la clasificación y el acopio de los datos.

Verificar:

- ✓ Revisar como están operando el algoritmo de búsqueda, o el script para la clasificación y el acopio de los datos, para datos no estructurados

Requisitos inherentes a la calidad de las bases de datos



Material de Curso Virtual Formación de Auditores Internos NTC PE 1000 (Julio 2021)

Requisito 2020:

7.3.3 La entidad debe documentar los manuales de uso y mantenimiento de los aplicativos.

Verificar:

- ✓ La documentación se encuentra actualizada para el uso y mantenimiento del software que se requiere para el desarrollo de la OE, de acuerdo con las especificaciones del diseño

Requisitos inherentes a la calidad de las bases de datos



Material de Curso Virtual Formación de Auditores Internos NTC PE 1000 (Julio 2021)

Requisito 2017:

7.4 PROCESAMIENTO

7.4.1 Integración de datos

La entidad debe documentar, implementar y mantener las actividades para la integración de datos.

Verificar:

- ✓ Verificar como se realiza la transmisión y consolidación de los archivos de datos

Requisitos inherentes a la calidad de las bases de datos

Material de Curso Virtual Formación de Auditores Internos NTC PE 1000 (Julio 2021)

Requisito 2017:

7.4 PROCESAMIENTO

7.4.2 Clasificación y codificación
La entidad debe documentar, implementar y garantizar que la clasificación y codificación de datos de acuerdo con lo señalado en el numeral 6.3.4. (nomenclaturas y clasificaciones)

Verificar:

- ✓ Importante establecer, que si se describe el uso de nomenclaturas y clasificaciones, estas estén representadas y contenidas en variables de los archivos de datos.

Requisitos inherentes a la calidad de las bases de datos

Material de Curso Virtual Formación de Auditores Internos NTC PE 1000 (Julio 2021)

Requisito 2017:

7.4 PROCESAMIENTO

7.4.3 Validación de la información recolectada

La entidad debe implementar y garantizar la aplicación de las reglas de validación y consistencia a la información recolectada.

Verificar:

- ✓ Importante establecer, si se disponen de indicadores o métricas de calidad de los datos para la información recolectada y procesada. (Inconsistencias encontradas, completitud etc.)

Requisitos inherentes a la calidad de las bases de datos



Material del Curso Virtual Formación de Auditores internos NTC PE 1000 (Julio 2021)

Requisito 2020:

7.7.2 Se deberían incluir pruebas unitarias, funcionales o integrales al software aplicativo para comprobar que las validaciones están operando de acuerdo con las especificaciones establecidas.

Verificar:

- ✓ Verificar que las pruebas sean consistentes con lo especificado en el numeral:
  - ✓ 6.11.2 La operación estadística debe considerar las pruebas funcionales y no funcionales para los aplicativos.

Material de Curso Virtual Formación de Auditores Internos NTC PE 1000 (Julio 2021)

Capitulo 8
Recolección o acopio
NTC PE 1000:2020

Requisitos inherentes a la calidad de las bases de datos



Material de Curso Virtual Formación de Auditores Internos NTC PE 1000 (Julio 2021)

Requisito 2020:

8.3.3 La entidad debe consolidar en un archivo los datos recolectados

Verificar:

- ✓ Revisar como se consolida la información recolectada en los archivos de datos y como se establecen los controles de esa consolidación.

Requisitos inherentes a la calidad de las bases de datos

Material de Curso Virtual Formación de Auditores Internos NTC PE 1000 (Julio 2021)

Requisito 2020:

8.4 La operación estadística debe finalizar los archivos de datos de acuerdo con lo establecido (véase 6.10).

Verificar:

- ✓ Revisar si se cumplieron los criterios establecidos para el cierre de los archivos de datos. Según el numeral 6.10

Material de Curso Virtual Formación de Auditores Internos NTC PE 1000 (Julio 2021)

Capitulo 9
Procesamiento
NTC PE 1000:2020

Requisitos inherentes a la calidad de las bases de datos



Material de Curso Virtual Formación de Auditores Internos NTC PE 1000 (Julio 2021)

Requisito 2020:

9. PROCESAMIENTO

La operación estadística debe realizar el procesamiento de los datos recolectados o acopiados de acuerdo con lo diseñado.

Verificar:

- ✓ Verificar integralmente que el procesamiento este acorde con el diseño, según el numeral 6.6

Requisitos inherentes a la calidad de las bases de datos

Material de Curso Virtual Formación de Auditores Internos NTC PE 1000 (Julio 2021)

Requisito 2020:

9.1 INTEGRACIÓN DE DATOS

La operación estadística debe implementar y conservar las actividades para la integración de datos.

Verificar:

- ✓ En caso de que se haga acopio de información de diferentes fuentes se revisara la integración de datos y los procesos técnicos y de negocio que se utilizan para combinar información de diferentes fuentes para convertirla en datos fiables y valiosos.

Requisitos inherentes a la calidad de las bases de datos



Material de Curso Virtual Formación de Auditores Internos NTC PE 1000 (Julio 2021)

Requisito 2020:

9.3.1 La operación estadística debe revisar los datos recolectados o acopiados, de acuerdo con lo diseñado.

Verificar:

- ✓ Se debe revisar como se está realizando la recolección de acuerdo con lo establecido en el numeral 6.5.1. Tales como métodos e instrumentos de recolección

Requisitos inherentes a la calidad de las bases de datos

Requisito 2020:

9.6 FINALIZACIÓN DE LOS ARCHIVOS DE DATOS

Los datos recolectados o acopiados deben almacenarse con medidas que aseguren, como mínimo:

a) la confidencialidad,
b) la disponibilidad,
c) el control de acceso,
d) la copia de seguridad,
e) las versiones con fecha y hora,
f) la protección de pérdida de información, y
g) la protección de modificación o alteración de acceso.

Verificar:

- ✓ Politica o lineamientos de backup de la entidad
- ✓ Almacenamiento de la base de datos final
- ✓ Perdida de información
- ✓ Log de verificación de cambios en los datos

Material de Curso Virtual Formación de Auditores Internos NTC PE 1000 (Julio 2021)