

**Departamento Administrativo
Nacional de Estadística**



**Producción Estadística
PES**

**Dirección de Metodología y Producción Estadística /
DIMPE**

**METODOLOGÍA GENERAL ENCUESTA DE TECNOLOGÍAS DE LA
INFORMACIÓN Y LAS COMUNICACIONES EN HOGARES –
ENTIC HOGARES**

Julio/2025

TABLA DE CONTENIDO

INTRODUCCIÓN	4
1. ANTECEDENTES	5
2. DISEÑO DE LA OPERACIÓN ESTADÍSTICA	8
2.1. DISEÑO TEMÁTICO	8
2.1.1. Necesidades de Información	8
2.1.2. Formulación de objetivos.....	10
2.1.3. Alcance	11
2.1.4. Marco de referencia.....	12
2.1.5. Definición de variables y construcción de indicadores estadísticos	21
2.1.6. Resultados estadísticos	24
2.1.7. Estándares estadísticos utilizados	25
2.1.8. Diseño del cuestionario	26
2.1.9. Normas, especificaciones o reglas de edición e imputación de datos	27
2.2. DISEÑO ESTADÍSTICO	27
2.2.1. Universo de estudio	28
2.2.2. Población objetivo	28
2.2.3. Cobertura geográfica.....	28
2.2.4. Desagregación geográfica	30
2.2.5. Desagregación temática.....	30
2.2.6. Fuentes de datos.....	30
2.2.7. Unidades estadísticas	30
2.2.8. Período de referencia	31
2.2.9. Periodo de recolección y frecuencia	31
2.2.10. Marco estadístico muestral.....	32
2.2.11. Diseño muestral	33
2.2.12. Ajustes de cobertura (o ajuste de cobertura por no respuesta).....	38
2.2.13. Especificaciones de ponderadores.....	46
2.3. DISEÑO DE LA RECOLECCIÓN/ ACOPIO	47

2.3.1. Métodos y estrategias de recolección de datos.....	47
2.3.2. Estructura organizacional del operativo y conformación del equipo	48
2.3.3. Esquema de entrenamiento de personal	50
2.3.4. Conformación del equipo	51
2.3.5. Proceso de sensibilización y acuerdos de intercambio	51
2.3.6. Elaboración de manuales	52
2.3.7. Diseño de las estrategias de comunicación y plan de contingencias	53
2.3.8. Diseño de la estrategia de seguimiento y control	53
2.3.9. Diseño de sistemas para la obtención de datos.....	57
2.3.10. Transmisión de datos	61
2.4 DISEÑO DE PROCESAMIENTO.....	62
2.4.1. Consolidación de archivos de datos	62
2.4.2. Codificación	64
2.4.3. Diccionario de datos.....	64
2.4.4. Revisión y validación.....	64
2.4.5. Diseño de instrumentos de edición (validación y consistencia) e imputación de datos	65
2.4.6. Diseño para la generación de cuadros de resultados.....	66
2.5. DISEÑO DEL ANÁLISIS.....	66
2.5.1. Métodos de análisis de resultados.....	67
2.5.2. Anonimización de microdatos.....	68
2.5.3. Verificación de la anonimización de microdatos.....	69
2.5.4. Comités de expertos.....	69
2.6. DISEÑO DE LA DIFUSIÓN Y COMUNICACIÓN.....	70
2.6.1. Diseño de sistemas de salida.....	70
2.6.2. Diseño de productos de difusión y comunicación	70
2.6.3. Entrega de productos	71
2.6.4. Estrategia de servicio	71
2.8. DISEÑO DE LOS SISTEMAS DE PRODUCCIÓN Y FLUJOS DE TRABAJO.....	73
3. DOCUMENTACIÓN RELACIONADA	75

INTRODUCCIÓN

Las encuestas elaboradas a los hogares constituyen una de las principales fuentes de datos socioeconómicos con las que cuentan los países, puesto que a partir de la información obtenida se calculan indicadores para la medición de variados aspectos económicos y sociales. Así mismo, se facilita el conocimiento y la explicación de determinantes o factores causales del comportamiento de la población objeto de estudio, lo cual es importante para el diseño, monitoreo y medición de resultados de políticas públicas.

El uso y la apropiación de las Tecnologías de la Información y las Comunicaciones (TIC) por parte de las personas y los hogares inciden en las preferencias, la productividad y la dinámica interna de la unidad familiar conformada por personas que acceden y se apropian de estas tecnologías. Algunos de los impactos sociales más importantes de acceso y uso de las TIC a través de los hogares se encuentran la oportunidad de acceder a información relevante, la búsqueda de empleo, la generación de ingresos, el intercambio de conocimiento, la creación de habilidades y capacidades y el logro de mayores niveles de integración social. De esta manera, las TIC inciden positivamente en la productividad y la calidad de vida de los hogares y personas.

La necesidad de monitorear las tendencias asociadas a este tipo de desarrollo es reconocida ampliamente por diferentes instituciones del país, pues a nivel de formulación y evaluación de políticas públicas y público-privadas, los planes y programas de estímulo y masificación social de las TIC para Colombia tienden a formularse actualmente con una mejor articulación técnica y mayor congruencia conceptual en relación con los indicadores sociales de acceso y uso de TIC producidos por el Departamento Administrativo Nacional de Estadística (DANE).

Como consecuencia de lo expuesto, a partir del 2007 y hasta 2011, el DANE calcula para Colombia los indicadores básicos de tenencia y uso de las Tecnologías de la Información y las Comunicaciones (TIC) en hogares y por personas, por medio de módulos insertados en la Gran Encuesta Integrada de Hogares (GEIH) y en la Encuesta de Calidad de Vida (ECV) a partir del 2012. Estos indicadores forman parte del programa internacional de medición de la Sociedad de la Información, entendida como el estado del desarrollo económico y social en el cual los individuos y agrupaciones acceden, se apropian, usan y adaptan las TIC de forma frecuente, intensiva, diversificada y significativa para sus vidas.

Para el 2018, Colombia inició el proceso de ingreso a la Organización para la Cooperación y el Desarrollo Económicos (OCDE) y definió un marco de trabajo para medir la transformación digital, por esta razón y con la necesidad de medir información especializada con componentes clave de la economía digital

y sus habilitadores (CONPES 3975 de 2019) surgió la Encuesta Tecnologías de la Información y la Comunicación en Hogares (ENTIC) como un instrumento que el DANE desarrolla con el objeto de medir y atender los nuevos requerimientos internacionales sobre medición de la sociedad de la información. Estos indicadores de tenencia y uso de TIC preservan un marco teórico acorde con los acuerdos determinados por expertos nacionales e internacionales acerca del diseño, aplicación e interpretación de encuestas nacionales de TIC y tiene en cuenta los lineamientos del Manual para la medición de acceso y uso de TIC en hogares y personas expedido por la Unión Internacional de Telecomunicaciones (UIT) en el 2020.

Ahora bien, el desarrollo de la ENTIC Hogares utiliza como período de referencia los últimos doce meses y generalmente el periodo de recolección se da entre los meses de agosto y octubre. El formulario de la ENTIC Hogares incluye cinco (5) módulos, con uno inicial de localización y vivienda. Finalmente, este documento contiene los aspectos metodológicos centrales de la ENTIC organizados en tres secciones, la primera sección corresponde a los antecedentes de la encuesta, la segunda sección contiene el diseño de la operación estadística desde la parte temática, estadística, de recolección/acopio, de procesamiento, de análisis, de difusión y declaración, así como la evaluación de las fases del proceso y de los sistemas de producción y flujos de trabajo y para concluir, la tercera sección hace referencia a la documentación relacionada, así mismo, se presenta el glosario, la bibliografía y los anexos.

1. ANTECEDENTES

La Encuesta Tecnologías de la Información y la Comunicación (ENTIC) Hogares, es consecuencia de la necesidad de medir indicadores especializados de acceso y uso de las TIC, así como de monitorear las tendencias asociadas a la transformación digital que se extiende a todos los sectores y aspectos de la sociedad. Dicha medición ha evolucionado y se ha especializado dentro del sistema de estadísticas nacional, adoptando estándares internacionales. Por su parte, en el contexto internacional, en el 2003 se creó el Observatorio para la Sociedad de la Información en Latinoamérica y el Caribe (OSILAC) como una estrategia de la Comisión Económica para América Latina y el Caribe (CEPAL) y del Instituto para la Conectividad en las Américas (ICA) adscrito al Centro Internacional de Investigación para el Desarrollo (CIID-IDRC). El objetivo de la presente Operación Estadística es el de promover la armonización estadística con el fin de fortalecer el monitoreo de las políticas y proyectos TIC en América Latina y el Caribe, así como apoyar a los países en la recopilación y análisis de datos estadísticos.

Para el 2004 y 2011, la OSILAC presenta un análisis de las estadísticas sobre Sociedad de la Información en los Institutos Nacionales de Estadística de América Latina y el Caribe. En este sentido, a través de la OSILAC, Colombia y la región participan en las siguientes iniciativas.

1. Partnership sobre medición de las TIC para el desarrollo: Busca definir un conjunto común de indicadores de TIC y ayudar a los países en desarrollo en sus esfuerzos para producir estadísticas armonizadas en esta área, contribuyendo a la disminución de la brecha de información que existe entre los países desarrollados y los no desarrollados (CEPAL, 2011).

2. Cumbre Mundial para la Sociedad de la Información (CMSI): Los distintos gobiernos aprobaron un plan de acción para promover la Sociedad de la Información en cada país, haciendo sugerencias para al desarrollo de indicadores que permitan evaluar el estado de las TIC y su avance, así como el monitoreo (benchmarking) de la implementación de dicho plan.

A partir del 2007, el DANE comenzó a calcular los indicadores básicos de tenencia y uso de las Tecnologías de la Información y las Comunicaciones (TIC) en hogares y personas, a través de un módulo insertado en la Gran Encuesta Integrada de Hogares (GEIH). Posteriormente, en septiembre de 2008, se publicaron los indicadores básicos con información del año 2007, dicho proceso se llevó a cabo hasta el 2011. Es importante mencionar que los indicadores que componen los boletines siguen las normativas internacionales de la Conferencia de las Naciones Unidas sobre Comercio y Desarrollo (UNCTAD/UN) y la CEPAL-OSILAC, para facilitar la comparación internacional de cifras. A partir del 2010 se añadió la referencia del partnership de la Unión Internacional de Telecomunicaciones (UIT), con algunas adaptaciones para Colombia.

En el 2012, se inició la recopilación de los indicadores a través de un módulo insertado en la Encuesta de Calidad de Vida (ECV), debido al creciente interés en el tema y a la decisión de monitorearlo mediante una operación estadística más enfocada en medir condiciones socioeconómicas. Por esta razón, el boletín de prensa publicado en 2013 presentó los resultados con base en la información básica sobre tenencia y uso de TIC recopilada en 2012. Este proceso se repite hasta la fecha acompañado de un documento de comunicado de prensa.

Por otra parte, la OCDE sostiene que para entender el desarrollo digital de un país es esencial realizar el seguimiento y análisis de indicadores y tendencias clave, razón por la cual, en virtud de la membresía de Colombia en la OCDE a partir de 2020, el país se ha comprometido a medir indicadores de transformación digital, proporcionando así una evaluación integral del estado del país en este ámbito.

Ahora bien, la hoja de ruta de medición "Going Digital" de la OCDE se presenta como una herramienta fundamental para alinear las áreas prioritarias de los países en la medición de la transformación digital mediante metodologías y enfoques comunes, la misma se desarrollada en colaboración con todos los organismos estadísticos relevantes de la OCDE en 2019 y actualizada

en 2022, esta hoja de ruta identifica diez acciones para mejorar la capacidad de los países en el monitoreo de la transformación digital y sus impactos. La hoja de ruta reconoce la necesidad de adaptar y expandir los sistemas estadísticos nacionales para reflejar adecuadamente la digitalización de nuestras economías y sociedades, incluyendo una perspectiva de género. Así mismo, se destaca la necesidad de nuevas infraestructuras de datos complementarias capaces de monitorear las actividades digitales y los flujos de datos de manera oportuna en cualquier lugar que ocurran.

En el Plan Nacional de Desarrollo (PND) 2018-2022 a través del “Pacto por la Transformación Digital de Colombia: Gobierno, empresas y hogares conectados con la era del conocimiento”, se determinó que las TIC habilitan la agregación de valor transversal en la economía, generan nuevos negocios y son la puerta de entrada a la industria 4.0. En este sentido, el PND plantea dos vías para promover la transformación digital. Primero, se hace referencia a la masificación de Internet de banda ancha e inclusión digital para la población colombiana, a través del fortalecimiento del marco normativo y segundo, a la implementación de las tecnologías digitales avanzadas, tales como el Blockchain, Internet de las Cosas¹ (IoT por sus siglas en inglés), entre otras, y así buscar más eficiencia en las relaciones entre mercados, ciudadanos y estado conforme a lo dispuesto en el documento CONPES 3975 de 2019.

Así mismo, a mediados de 2018 y durante 2019 se desarrollaron mesas de trabajo con personal técnico del Ministerio de Tecnologías de la Información y Comunicaciones (MINTIC), con el objetivo de diseñar una gran encuesta que no sólo midiera los indicadores básicos sobre acceso y uso de TIC que actualmente publica el DANE, sino que permitiera analizar y atender nuevos requerimientos internacionales sobre medición de la sociedad de la información y generar indicadores especializados que permitan promover el proceso de transformación digital en Colombia. Así mismo, el documento CONPES 3975 de 2019, en su línea de acción 3, indica que el MINTIC y el DANE deberán aplicar la encuesta a nivel nacional con un nivel de desagregación departamental para hogares.

Actualmente el DANE cuenta con información de indicadores básicos sobre acceso y uso de las TIC por las personas, hogares y empresas, y la realización de dos encuestas especializadas sobre tenencia y uso de TIC en hogares ENTIC Hogares para los periodos de referencia 2020 y 2021, como resultado de los convenios interadministrativos 781 de 2019 y 899 de 2021 suscritos entre el DANE y el MINTIC.

¹ Internet de las cosas: Internet of Things (IoT): Dispositivos interconectados que recolectan e intercambian datos y pueden ser monitoreados o controlados remotamente mediante Internet, sin requerir la interacción entre humanos o entre humanos y computadores. Se excluyen los sensores y los dispositivos de detección simple de movimiento, sonido, temperatura, humo, etc. (Internet of things (IoT). Obtenido de TechTarget: (febrero de 2020)). Ver conceptos.

Por su parte, el Plan Nacional de Desarrollo 2022-2026 "Colombia Potencia Mundial de la Vida" tiene como objetivo establecer las bases para que el país se posicione como líder en la protección de la vida, mediante la instauración de un nuevo contrato social que propicie la superación de injusticias históricas, la prevención de conflictos, la modificación de la interacción con el entorno y una transformación productiva respaldada en el conocimiento y en armonía con la naturaleza.

Dentro de este contexto, el sector de Tecnologías de la Información adquiere relevancia en la denominada "Seguridad humana y justicia social" de igual forma, la Ley 2294 de 2023, que expide el PND, establece que el acceso y uso de las tecnologías digitales deben considerarse un derecho fundamental en lugar de un privilegio. Así mismo, enfatiza que el Estado tiene la responsabilidad de garantizar la conexión y alfabetización digital de los ciudadanos, impulsando la sociedad del conocimiento a través de las Tecnologías de la Información y las Comunicaciones (TIC) con un enfoque diferencial, buscando generar oportunidades, equidad, riqueza y productividad. (DNP, 2023, p. 546).

2. DISEÑO DE LA OPERACIÓN ESTADÍSTICA

En este capítulo se presenta el diseño metodológico de la ENTIC Hogares, cuyo diseño determina los métodos y procedimientos con los que se construyen los instrumentos, procesos y actividades, se recolectan, procesarán y analizarán los datos, se difunde la información y evalúan las fases del proceso estadístico para cumplir los objetivos y satisfacer las necesidades de información.

2.1. DISEÑO TEMÁTICO

Este componente comprende aspectos tales como las necesidades de información de los usuarios, los objetivos generales y específicos de la ENTIC Hogares, los diferentes marcos que se considera necesario incluir, el diseño de indicadores y del formulario, así como las nomenclaturas utilizadas para entender la planeación de la operación estadística.

2.1.1. Necesidades de Información

Tal y como se mencionó previamente, la ENTIC Hogares surgió de la necesidad de medir indicadores especializados de acceso y uso de las TIC, además de monitorear las tendencias asociadas a la transformación digital que se extiende a la totalidad los sectores y aspectos de la sociedad. Así mismo, el país avanza en la armonización estadística frente a indicadores TIC dentro del contexto internacional, tomando como referencia las normativas internacionales de la Conferencia de las Naciones Unidas sobre Comercio y Desarrollo (UNCTAD/UN), la CEPAL-OSILAC y la Unión Internacional de Telecomunicaciones (UIT), con algunas adaptaciones para Colombia.

El Plan Nacional de Desarrollo (PND) 2022-2026 propone una estrategia de datos sectorial en línea con las recomendaciones y directrices emitidas por el Comité Nacional de Datos y el Comité de Administración de Datos. El propósito es incrementar la disponibilidad y aprovechamiento de datos de calidad, impulsar proyectos de uso e intercambio de datos y consolidar una cultura de datos en los sectores administrativos del país.

El documento también destaca que la disponibilidad de datos de calidad es un elemento esencial para la formulación de políticas públicas de inclusión y protección social. (DNP, pág. 74). Por su parte, la Ley 2294 de 2023 dicta medidas específicas para respaldar estas prioridades, incluyendo el fomento de la conectividad digital como generador de oportunidades, la implementación de estrategias integrales para democratizar las TIC y el fortalecimiento del sector TIC como motor de crecimiento, empleo y desarrollo nacional, así:

Artículo 142. Conectividad digital para cambiar vidas: *“Para efectos de promover la conectividad digital como un generador de oportunidades, riqueza, igualdad y productividad”* (Congreso de la República, 2023).

Artículo 143. Transformación digital como motor de oportunidades e igualdad: *“El Ministerio de Tecnologías de la Información y las Comunicaciones diseñará e implementará una estrategia integral para democratizar las TIC y desarrollar la sociedad del conocimiento y la tecnología en el país”* (Congreso de la República, 2023).

Artículo 144. Fortalecimiento del sector TIC: *“El Ministerio de Tecnologías de la Información y las Comunicaciones promoverá la consolidación de la Industria TIC nacional como un motor de crecimiento, empleo y desarrollo para el país”* (Congreso de la República, 2023).

En concordancia con lo anterior y reconociendo la importancia de los datos de calidad, el Plan Nacional de Desarrollo (PND) propone una estrategia de datos sectoriales en relación con las recomendaciones y directrices emitidas por el Comité Nacional de Datos y el Comité de Administración de Datos. El objetivo es incrementar la disponibilidad y aprovechamiento de datos de calidad, impulsar proyectos de uso e intercambio de datos, y consolidar una cultura de datos en los sectores administrativos del país.

Ahora bien, la Organización para la Cooperación y el Desarrollo Económicos (OCDE), la cual es integrada por un total de 38 países miembros, los cuales forman una comunidad de naciones comprometidas con los valores que giran en torno a la democracia basada en el estado de derecho, los derechos humanos y una economía de mercado abierta y transparente dentro de los cuales se encuentra Colombia, sostiene que para entender el desarrollo digital de un país es esencial realizar el seguimiento y análisis de indicadores y tendencias clave. En virtud de la membresía de Colombia

en la OCDE, el país se ha comprometido a medir indicadores de transformación digital, proporcionando así una evaluación integral del estado del país en este ámbito.

La apropiación de las Tecnologías de la Información y la Comunicación (TIC) por parte de las personas y los hogares inciden en las preferencias, la productividad y la dinámica interna de la unidad familiar conformada por personas que acceden y se apropian de estas tecnologías. Entre los impactos sociales más importantes de acceso y uso de las TIC a través de los hogares, se cuentan la mayor oportunidad de acceder a información relevante a la búsqueda de empleo y la generación de ingresos, el intercambio de conocimientos, la creación de habilidades y capacidades y promover mayores niveles de integración social. De esta manera, las TIC pueden incidir positivamente en la productividad y la calidad de vida de los hogares y personas.

Con los indicadores especializados de la ENTIC, se facilitarán las comparaciones nacionales e internacionales en el sector TIC y será un insumo de información estadística para la formulación, seguimiento y evaluación de las políticas públicas sectoriales.

Esta identificación de las necesidades de información, se generan a través de las mesas de trabajo interinstitucionales, los comités internos y externos que se realizan para cada publicación, y el contacto con los usuarios académicos a través de las solicitudes realizadas por la Sala de Procesamiento Especializado, cuyas necesidades son analizadas y priorizadas por las mesas interinstitucionales.

La identificación de necesidades de información estadística se constituye en el elemento base para la producción estadística. El Grupo interno de trabajo de Industria, atiende el procedimiento institucional para la detección y análisis de necesidades que inicia con la identificación y la consolidación de las principales necesidades expuestas por los grupos de interés. Al final de cada mes se consolidan las necesidades de información estadística que los usuarios de las Estadísticas de Industria reportan al DANE, estas surten el proceso de confirmación con los usuarios. Posteriormente de manera trimestral e interna se realiza: la priorización, el análisis de viabilidad técnica y económica, y se elabora un informe de priorización de necesidades de información estadística (el cual será publicado a partir de julio de 2022 en la página web de la entidad).

2.1.2. Formulación de objetivos

Objetivo general:

Obtener información que permita analizar y realizar comparaciones de las condiciones de tenencia y uso de las TIC en hogares colombianos, las cuales posibiliten el seguimiento y la formulación de políticas enfocadas al desarrollo y aprovechamiento de las TIC, así como hacer seguimiento a las variables necesarias para el diseño e implementación de políticas públicas y público-privadas, los planes y programas de estímulo y masificación social de las TIC para Colombia, con mayor articulación técnica y congruencia conceptual.

Objetivos específicos:

- Obtener la información necesaria para la actualización de los indicadores de tenencia y uso de las TIC a nivel de viviendas, hogares y personas, cuyos resultados sirvan como insumo para la definición de políticas que permitan diseñar y ejecutar planes sociales.
- Brindar información que permita la obtención de los indicadores para determinar avances en la Sociedad de la Información.
- Obtener información que posibilite estandarizar y facilitar la comparación internacional de las cifras en temáticas TIC.

2.1.3. Alcance

La ENTIC Hogares es una investigación que permite obtener información especializada sobre la tenencia y el uso de las TIC. La operación indaga por las características y tenencia de TIC a nivel de hogares y el uso de TIC a nivel de personas (que conforman los hogares).

En el caso de los **hogares**, la ENTIC consulta, entre otros, los siguientes aspectos:

- Tenencia de bienes y servicios TIC.
- Percepción de las condiciones de los servicios TIC.

Acerca de las **personas que componen los hogares**, el alcance de la ENTIC incluye las siguientes dimensiones:

- Aspectos de caracterización demográfica (edad, sexo, parentesco con el jefe del hogar, estado civil, educación, trabajo).
- Uso de Tecnologías de la Información y la Comunicación - TIC.
- Seguridad y uso responsable de internet.

La ENTIC Hogares tiene representatividad para el total nacional, departamental y en relación con sus áreas (cabecera - centros poblados y rural disperso), lo cual se logra por medio de una muestra aproximada de 53.300 hogares con entrevista directa a las personas miembros del hogar con edad igual o superior a 5 años seleccionados en la muestra. De igual forma, en la sección donde se describe el diseño del formulario se presentan con más detalle cada uno de los aspectos abordados.

2.1.4. Marco de referencia

El marco de referencia de la ENTIC Hogares, incluye los principales antecedentes teóricos de la medición y análisis de los procesos de transformación digital de la sociedad y lo que se conoce como la cuarta revolución industrial, así como los enfoques más representativos para su abordaje, algunos conceptos fundamentales relacionados con la infraestructura TIC y los referentes internacionales más relevantes en el tema.

La revisión a continuación evidencia el impacto de los procesos de TIC sobre la sociedad y la necesidad de medir los procesos de transformación digital y su composición en dimensiones objetivas y subjetivas, aspectos que son considerados e incorporados en la Encuesta Tecnologías de la Información y la Comunicación en hogares del DANE.

Marco teórico

Para la CEPAL² la Declaración del Milenio de Naciones Unidas en el 2000, la cual fue firmada por 189 países en Nueva York, fue uno de los primeros escenarios en los cuales se manifestó la necesidad de disponer de un cuerpo de indicadores TIC dentro del monitoreo del desarrollo socioeconómico de los países. En este contexto, se consideraba importante realizar el seguimiento al número de líneas de teléfono y teléfonos celulares, así como al uso de computadores personales y número de usuarios de internet. Posteriormente, la Cumbre Mundial para la Sociedad de la Información (CMSI), desarrollada en el 2003 en Suiza y en el 2005 en Túnez, fue la base para promover la implementación de un plan de acción para los procesos de medición de acceso y uso de las TIC, de modo que para lograrlo se estableció el Partnership en medición TIC para el desarrollo.

Posteriormente, en el 2005, este Partnership publicó una lista clave de indicadores para medición de TIC con el objetivo de mejorar la comparabilidad internacional. En principio, distingue indicadores básicos para los países donde el uso de las tecnologías va en aumento y que posean un interés en observar el proceso de evolución en TIC, así como indicadores extendidos para países con altos niveles en la utilización de tecnologías y un sector creciente de TIC (UIT, 2005). Esta primera aproximación contaba con cuatro series de indicadores: (i) infraestructura de las TIC y acceso a ellas; (ii) acceso a las TIC y su uso por parte de hogares e individuos; (iii) uso de las TIC por las empresas, y (iv) sector de las TIC y comercio de bienes relacionados con ellas. Estos indicadores, fueron actualizados en el 2010 con una lista revisada que contiene 46 indicadores clave sobre TIC y dos indicadores de referencia, comparada con la lista del 2005 que contenía 41 indicadores TIC y un indicador de referencia (UIT, 2010).

A partir del 2008, el Partnership ha publicado en su página web (de la UIT) reportes estadísticos con base en los indicadores. Además, en el 2009 se publicaron dos manuales para la aplicación de estos

² Comisión Económica para América Latina y el Caribe (CEPAL). (2011). Caracterización del proceso de creación de estadísticas armonizadas sobre tecnologías de la información y la comunicación en los organismos nacionales de estadísticas.

indicadores, los cuales son ampliamente utilizados por los países e Institutos Nacionales de Estadísticas (INE). Para la temática de hogares se publicó el Manual para la medición del uso y el acceso a las TIC por los hogares y las personas de la UIT (2009) y, en el caso de las empresas, el Manual para la producción de estadísticas sobre la economía de la información de la UNCTAD (2009). Inclusive en el 2009, la UNESCO publicó un nuevo manual con el objetivo de enriquecer el conjunto básico de indicadores TIC para el tema específico de educación y plantear una lista ampliada de indicadores o mediciones indirectas (proxy), que permitan abordar un espectro más amplio de temas de política educativa.

Para el año 2020, la UIT³ realizó un proceso de actualización del “Manual para la medición del acceso y el uso de las TIC en los hogares y por las personas”, donde establece que, para la recopilación de estadísticas sobre las TIC a nivel de los hogares, existe la necesidad de ajuste para los métodos y conceptos estadísticos, con el objetivo de estar al día en los cambios que experimenta el área de las tecnologías de la información y la comunicación.

Los anteriores esfuerzos realizados por las organizaciones mundiales y los países surgen de reconocer a las TIC como factor facilitador del desarrollo al ser estas aplicadas y utilizadas de la manera adecuada. Es crítico para los países que se dirigen hacia la creación de sociedades de la información, el aprovechamiento de las TIC y el reconocimiento de la dimensión social de las repercusiones de estas, incluyendo aspectos relativos a la brecha digital (desigualdad de acceso y de utilización de las TIC). También la integración social a través de las TIC, así como las características de acceso a las tecnologías y su uso por distintos grupos demográficos y sociales (UIT, 2014). Con la disponibilidad de datos y la comparabilidad entre los países, los responsables políticos pueden tomar decisiones informadas y explotar las ventajas de las TIC para sus ciudadanos. En conclusión, para la UIT (2014) el crecimiento económico y el desarrollo no explotarán su máximo potencial si las economías no son capaces de hacer un uso adecuado de las nuevas tecnologías.

El acelerado proceso de transformación digital y el reconocimiento de las TIC como motor de desarrollo económico y social implican nuevos retos en términos de información e impulsan la necesidad de producir estadísticas fiables, completas y comparables para ayudar a los gobiernos y al sector TIC, en la implementación de políticas públicas (UIT, 2014). Lo anterior representa, según la OCDE (2019) dos retos. El primero de corto plazo, establece la necesidad de mejorar la comparabilidad internacional de los indicadores vigentes y flexibilizar los sistemas estadísticos para que respondan de mejor manera a los nuevos conceptos y de rápida evolución, impulsados por la transformación digital (OCDE, 2019); el segundo reto, de largo plazo, consiste en diseñar de forma interdisciplinaria nuevos enfoques para recopilar y usar la información capturada con sistemas digitales, toda vez que la transformación digital se extiende a todos los sectores y aspectos de la sociedad (OCDE, 2019), es importante medir la tenencia y el uso de las TIC en los hogares y por

³ Unión Internacional de Telecomunicaciones (UIT) (2020). Manual para la medición del uso y el acceso a las TIC por los hogares y las personas. Ginebra – Suiza.

personas, para lograr corroborar los progresos que han alcanzado los países en sus esfuerzos por convertirse en sociedades de la información (UIT, 2014).

La Organización para la Cooperación y el Desarrollo Económicos (OCDE)⁴ sostiene que para entender el desarrollo digital de un país es esencial realizar el seguimiento y análisis de indicadores y tendencias clave. En virtud de la membresía de Colombia en la OCDE⁵, se ha comprometido a medir indicadores de transformación digital, proporcionando así una evaluación integral del estado del país en este ámbito.

La hoja de ruta de medición "Going Digital" de la OCDE⁶ se presenta como una herramienta fundamental para alinear las áreas prioritarias de los países en la medición de la transformación digital mediante metodologías y enfoques comunes. Desarrollada en colaboración con todos los organismos estadísticos relevantes de la OCDE en 2019 y actualizada en 2022, dicha hoja de ruta identifica diez acciones para mejorar la capacidad de los países en el monitoreo de la transformación digital y sus impactos, la misma reconoce la necesidad de adaptar y expandir los sistemas estadísticos nacionales para reflejar adecuadamente la digitalización de nuestras economías y sociedades, incluyendo una perspectiva de género. Así mismo, se destaca la necesidad de nuevas infraestructuras de datos complementarias capaces de monitorear las actividades digitales y los flujos de datos de manera oportuna en cualquier lugar que ocurran.

Acciones propuestas por la OCDE:

- Visibilizar la economía digital en las estadísticas económicas.
- Comprender los impactos económicos de la transformación digital.
- Fomentar la medición de los impactos de la transformación digital en los objetivos sociales y el bienestar de las personas.
- Diseñar enfoques nuevos e interdisciplinarios para la recopilación de datos.
- Supervisar las tecnologías que sustentan la transformación digital, en particular Internet de las Cosas, Inteligencia Artificial y Blockchain.
- Mejorar la medición de datos y flujos de datos.
- Definir y medir las necesidades de habilidades para la transformación digital.
- Medir la confianza en los entornos en línea.
- Establecer un marco de evaluación de impacto para los gobiernos digitales.
- Ampliar la recopilación y accesibilidad de las estadísticas de género.

Además, la OCDE dispone del informe "Going Digital Toolkit", el cual asiste a los países en la evaluación de su estado de desarrollo digital y en la formulación de políticas mediante 44 indicadores organizados en siete dimensiones de política. El objetivo de esta herramienta es permitir

⁴ Organización para la Cooperación y el Desarrollo Económicos (OCDE). (2015). Perspectivas de la OCDE sobre la economía digital.

⁵ Organización para la Cooperación y el Desarrollo Económicos (OCDE). (2019). Cómo medir la transformación digital: Hoja de ruta para el futuro.

⁶ Organización para la Cooperación y el Desarrollo Económicos (OCDE). (2019). Going Digital: Shaping Policies, Improving Lives.

a los Estados formular estrategias de política y acciones como respuesta, facilitando evaluaciones comparativas con los estándares de la OCDE, la Unión Europea (EU28) o entre países, permitiendo realizar evaluaciones comparativas.

El texto de Sánchez-Torres, Gonzales y Sánchez-Muñoz (2012) a partir de una revisión de la literatura, resume las oportunidades que ofrecen las TIC para países en desarrollo. En primer lugar, identifica el incremento de la capacidad de los investigadores, científicos, profesionales, instituciones y gobiernos, entre otros, a través del acceso a la información y materiales que se consiguen con el desarrollo tecnológico. En segundo lugar, permiten que grupos excluidos y minoritarios puedan crear ciertas comunidades con capacidad de presión política y así promover la integración social. En tercer lugar, indica que favorecen los procesos de descentralización y permiten un mayor acercamiento de los gobiernos a los ciudadanos (Sánchez-Torres, Gonzales y Sánchez Muñoz, 2012).

Marco conceptual

De acuerdo con la Ley 1341 de 2009, Colombia define las TIC como el conjunto de recursos, herramientas, equipos, programas informáticos, aplicaciones, redes y medios que permiten la compilación, procesamiento, almacenamiento, transmisión de información como voz, datos, texto, video e imágenes. Además, se define la importancia del acceso a las TIC en términos de la sociedad de la información, pues en el artículo 3 se indica que *“El Estado reconoce que el acceso y uso de las Tecnologías de la Información y las Comunicaciones, el despliegue y uso eficiente de la infraestructura, el desarrollo de contenidos y aplicaciones, la protección a los usuarios, la formación de talento humano en estas tecnologías y su carácter transversal, son pilares para la consolidación de las sociedades de la información y del conocimiento”*(Congreso de la República, 2009).

Para Sánchez-Torres, Gonzáles y Sánchez-Muñoz (2012) las TIC son aquellas tecnologías que permiten la adquisición, almacenamiento, procesamiento, evaluación, transmisión, distribución y difusión de la información. Además, establece que algunos de los rasgos relevantes para las TIC son:

1. Tecnologías para actuar sobre la información, no solo información para actuar sobre la tecnología.
2. Tienen gran capacidad de penetración y efecto en la economía.
3. Cuentan con la capacidad y la lógica de interconexión que todo sistema de TI utiliza, permitiendo una interconexión rápida y barata.
4. Permiten reprogramar y reequipar a las organizaciones.
5. Convergencia de tecnologías específicas para el desarrollo de estas.

Este carácter transversal de las TIC sobre la vida humana hace que surja una concepción en la literatura de la evolución de la sociedad industrial hacia una “sociedad de la información”. Para la Unión Internacional de Comunicaciones (UIT, 2014) desde un punto de vista conceptual, la sociedad

de la información es un conjunto complejo de temas, entidades, acciones y relaciones. Para la Unión Europea, la sociedad de la información consiste en economías con un formidable potencial de crecimiento de empleo e integración donde *“Las tecnologías digitales hacen que cada vez sea más fácil y barato el acceso, tratamiento, almacenamiento y la transmisión de la información. La base de la nueva economía es la transformación de la información digital en valor económico y social, creando nuevas industrias modificando otras y afectando profundamente la vida de los ciudadanos”* (Sánchez et al, 2012).

Ahora bien, las tendencias que dirigen una economía del conocimiento o “sociedad del conocimiento”, se relacionan con la creciente codificación, conocimiento y desarrollo de nuevas tecnologías, el aumento en tasas de investigación e innovación, la creciente importancia de la educación y mayores ganancias en la cadena de valor debido a inversiones en intangibles, entre otros (Sánchez et al, 2012).

Es así, que conceptos como el acceso y el uso de las TIC se convierten en parte fundamental de las discusiones. En relación con dichos términos, la OCDE define la “Brecha Digital” como la brecha existente entre individuos, empresas y áreas geográficas de diferentes niveles socioeconómicos, con respecto a las oportunidades de acceder y usar las TIC. Es importante distinguir, tanto a nivel conceptual como para los procesos de medición, lo que se entiende por acceso a las TIC y por utilización de estas. El acceso a las TIC se refiere a la disponibilidad de estas tecnologías en los hogares, mientras que la utilización consiste en un uso de estas por uno o más miembros de una familia, ya sea en el hogar o en cualquier otro sitio (ITU, 2014).

Por último, para la OCDE (2015) los fundamentos de la economía digital son la internet, las redes de banda ancha, las aplicaciones móviles, los servicios de TI y el hardware. Esta Organización enmarca el proceso de la transformación digital en la digitalización, entendida como la conversión de una señal analógica que transporta información (por ejemplo, el sonido, la imagen y el texto impreso) a formas más eficientes y universales de recopilación de información (bits binarios). Estos datos digitales, luego de un proceso de recolección y digitalización, se pueden utilizar ilimitadamente por dispositivos digitales sin degradación, a velocidades muy altas y con un costo marginal insignificante (OCDE, 2015). El otro aspecto mencionado por la OCDE son las interconexiones crecientes, donde gracias al Internet permiten que esto se produzca de forma global.

Marco Legal o normativo

La Ley 1341 de 2009 marcó un referente histórico en el país, toda vez que promovió las Tecnologías de la Información y las Comunicaciones como servicio universal, en desarrollo de los artículos 16, 20 y 67 de la Constitución Política, los cuales determinan que el Estado debe propiciar que todo colombiano tenga acceso a las tecnologías de la información y las comunicaciones para el ejercicio pleno de derechos como la libre expresión, el libre desarrollo de la personalidad, el de informar y recibir información veraz e imparcial, la educación y el acceso al conocimiento.

De igual forma, a partir del Decreto 1078 de 2015 se recopiló la normatividad previa del Sector de Tecnologías de la Información y las Comunicaciones, donde se establecen los objetivos y funciones a cargo de MinTIC y los principios y conceptos sobre la sociedad de la información y la organización de las TIC. Posteriormente, el Plan Nacional de Desarrollo (PND) 2018-2022 a través del "Pacto por la Transformación Digital de Colombia: Gobierno, empresas y hogares conectados con la era del conocimiento" es el marco por medio del cual se plantea la estrategia para que las TIC sean el motor que dinamice nuevos modelos de negocio que conforman la industria 4.0.

Esta estrategia fue puesta en consideración del documento CONPES 3975 de 2019 que plantea la Política Nacional para la Transformación Digital por medio de la cual se establece la estrategia para cerrar brechas a nivel público y privado al incorporar tecnologías digitales que permita enfrentar los cambios económicos y sociales que representa la industria 4.0 y la inteligencia artificial.

Consecuentemente, se formula la Ley 2108 de 2021 la cual tiene por objeto establecer el internet como servicio público y de acceso esencial y universal. De esta forma, se garantiza que la prestación del servicio se realice de manera eficiente, continua y permanente, extendiendo este servicio a la totalidad de la población del país y priorizando aquella en condición de vulnerabilidad por condición social o étnica o por habitar en zonas rurales o apartadas.

En la actualidad, el Plan Nacional de Desarrollo 2022-2026 "Colombia Potencia Mundial de la Vida" tiene como finalidad establecer las bases para que el país se posicione como líder en la protección de la vida, mediante la instauración de un nuevo contrato social que propicie la superación de injusticias históricas, la prevención de conflictos, la modificación de la interacción con el entorno y una transformación productiva respaldada en el conocimiento y en armonía con la naturaleza, de modo que el sector de Tecnologías de la Información adquiere una destacada relevancia en la denominada "Seguridad humana y justicia social".

Así mismo, la Ley 2294 de 2023, la cual expide el Plan Nacional de Desarrollo, establece que el acceso y uso de las tecnologías digitales deben considerarse un derecho fundamental en lugar de un privilegio. Así mismo, enfatiza que el Estado tiene la responsabilidad de garantizar la conexión y alfabetización digital de los ciudadanos, impulsando la sociedad del conocimiento a través de las Tecnologías de la Información y las Comunicaciones (TIC) con un enfoque diferencial, buscando generar oportunidades, equidad, riqueza y productividad (DNP, 2023, p. 546)⁷.

Así mismo, dicho apartado normativo dicta medidas específicas para respaldar estas prioridades, incluyendo el fomento de la conectividad digital como generador de oportunidades, la implementación de estrategias integrales para democratizar las TIC y el fortalecimiento del sector TIC como motor de crecimiento, empleo y desarrollo nacional.

⁷ Departamento Nacional de Planeación (2023). Ley 2294 del 9 de mayo de 2023 "Por el cual se expide el Plan Nacional de Desarrollo 2022- 2026 "Colombia Potencia Mundial de la Vida". https://colaboracion.dnp.gov.co/CDT/portalDNP/PND-2023/Ley_2294_del_19_de_mayo_de_2023.pdf

Finalmente, la Ley 2335 de 2023 regula la producción de estadísticas oficiales, reconoce en el Sistema Estadístico Nacional un sistema de principios, instancias y sujetos que son responsables de la producción y difusión estadística. Entiende que las estadísticas oficiales permiten conocer las realidades del país y, así, la toma de decisiones. La importancia de la operación también se enmarca en el desarrollo de los principios estadísticos, entre estos, el de la confidencialidad de la información estadística de privados: Es una obligación de carácter legal del Departamento Administrativo Nacional de Estadística (DANE), preservar y garantizar la reserva de la información recolectada a través de censos, encuestas, operaciones estadísticas, estadísticas experimentales y registros administrativos de manera que no sea posible deducir la información de carácter individual a partir de los informes estadísticos presentados al público, ni su utilización con fines distintos a los estadísticos. (Ley 2335 de 2023, artículo 4, parágrafo 1). En el mismo sentido el parágrafo 2, indica: El DANE no podrá proveer ni divulgar información de carácter reservado o privado de las personas. Para el tratamiento de información de carácter sensible, el DANE garantizará lo establecido en la Ley 1581 de 2012, así como la reserva estadística contenida en la presente ley y la aplicación de altos estándares de calidad y buenas prácticas internacionales en la materia.

En este sentido la ENTIC-H debe garantizar la confidencialidad, entre otros principios, como elemento que facilita la calidad y veracidad de la información suministrada por las fuentes; factor estratégico para el éxito de la operación estadística.

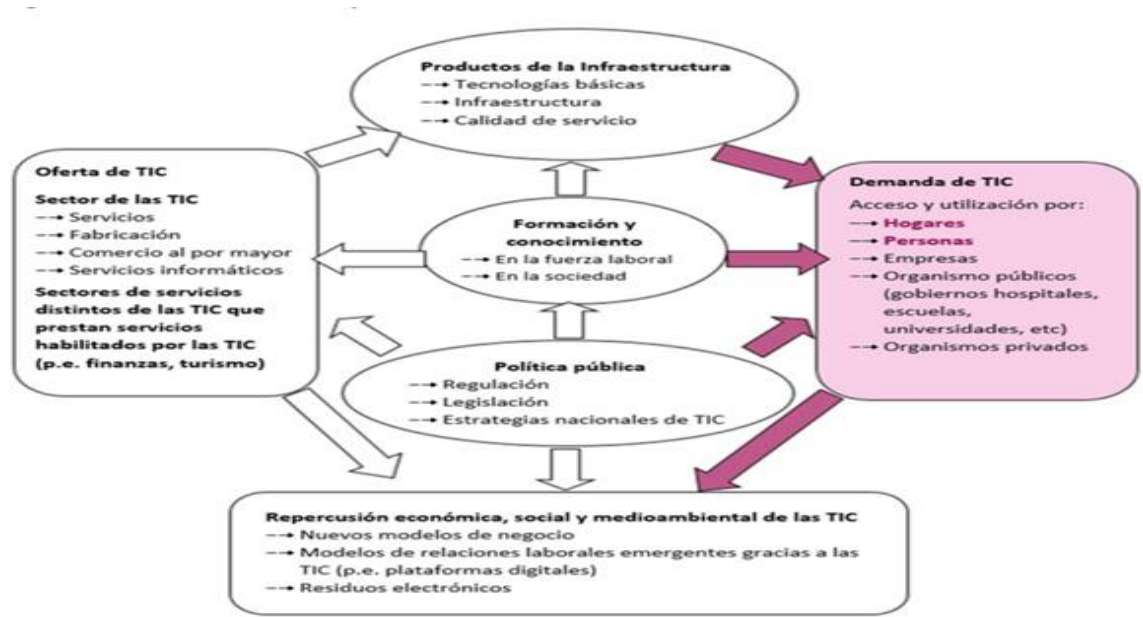
Referentes internacionales

Como se mencionó anteriormente, uno de los referentes internacionales más importantes es la Cumbre Mundial para la Sociedad de la Información (CMSI), realizada en el año 2003 en Suiza y en el 2005 en Túnez, la cual propuso un plan de acción para los procesos de medición de acceso y uso de las TIC. Para lograr lo anterior, se estableció el Partnership en medición de las TIC para el desarrollo, liderado por la Unión Internacional de Telecomunicaciones (UIT) y con participación de la Organización para la Cooperación y el Desarrollo Económico (OCDE), la Conferencia de las Naciones Unidas sobre Comercio y Desarrollo (UNCTAD), el Instituto de Estadísticas de la UNESCO (UIS), la Comisión Económica para América Latina y el Caribe (CEPAL-ECLAC), la Comisión Económica para África (CEPA), la Comisión Económica y Social para Asia y el Pacífico (CESPAP), la Comisión Económica y Social para Asia Occidental (CESPAO), la Task Force en TIC de Naciones Unidas y el Banco Mundial (BM).

En este contexto, de acuerdo con el Ministerio de Tecnologías de la Información y las Comunicaciones (MINTIC) la sociedad de la información es *“aquella en la cual las tecnologías que facilitan la creación, distribución y manipulación de la información juegan un papel importante en las actividades sociales, culturales y económicas debe estar centrada en la persona, integradora y orientada al desarrollo, en que todos puedan crear, consultar, utilizar y compartir la información y el conocimiento”* (Ministerio de Tecnologías de la Información y las Comunicaciones, S.F.). A

continuación, se presenta el marco conceptual de la Sociedad de la Información extraído del documento de la UIT (2020) donde se identifica la interacción entre oferta y demanda de las TIC y cómo se conectan a través de productos de la infraestructura, formación y conocimiento, política pública y como todo tiene una repercusión económica, social y medioambiental.

Figura 1. Marco conceptual de la Sociedad de la Información



Fuente: Unión Internacional de Telecomunicaciones, 2020.

En relación con lo anterior, por el lado de la oferta se proporcionan los bienes y servicios TIC. Es fundamental precisar que este tipo de infraestructura y sector se encuentra en constante evolución, cambiando inclusive su estructura y agregando de forma constante nuevas tecnologías, conceptos y plataformas; por ejemplo, con el desarrollo del internet se pasa al concepto del Internet de las Cosas (IoT por sus siglas en inglés) y la computación en la nube, entre otros. Lo anterior, afecta todas las dimensiones del ser humano y su comunicación, incluyendo otros sectores como la educación, el turismo, las finanzas, entre otros. (UIT,2020).

En este orden de ideas, se obliga a contar con una actualización constante en términos de conceptos y apropiación de tecnologías que deben ser incluidas en esta operación estadística, así como la integración de indicadores de interés en materia de política pública como la desigualdad en el acceso y utilización de las TIC por distintos grupos demográficos y sociales, entre otras variables de interés (UIT,2020).

Adicionalmente, se utiliza como referencia, la creación del Observatorio para la Sociedad de la Información en Latinoamérica y el Caribe (OSILAC) en el 2003, el cual tenía como objetivo, promover la armonización estadística, así como apoyar a los países en la recopilación y análisis de datos TIC en América Latina y el Caribe.

Para el 2004 y luego actualizado en 2011 la OSILAC presentó la “Caracterización del proceso de creación de estadísticas armonizadas sobre tecnologías de la información y la comunicación en los organismos nacionales de estadística”, el cual fue un documento por medio del cual se recopiló la información de 42 países de América Latina y el Caribe, determinando la demanda de los países en materia de recolección de información TIC y evaluando los avances sobre la medición de las TIC (CEPAL, 2011).

Siguiendo estos referentes, la ENTIC Hogares toma como referencia el manual para la medición de acceso y uso de TIC en hogares de la ITU y los documentos de la OCDE, que como lo indican, se encuentran encaminados a mejorar la comparabilidad internacional de los indicadores vigentes y flexibilizar los sistemas estadísticos para que respondan mejor a los nuevos conceptos y de rápida evolución impulsada por la transformación digital, así como crear nuevos enfoques para recopilar y usar la información capturada con sistemas digitales. Como consecuencia de lo anterior, se encuentran referencias como el Going Digital: Shaping Policies, Improving Lives y Cómo medir la transformación digital: Hoja de ruta para el futuro publicados por la OCDE en 2019.

Referentes nacionales

Debido a la temática específica de esta operación estadística, no existen ejercicios similares que midan las dimensiones de acceso y uso de TIC en Colombia con este nivel de detalle; sin embargo, tal y como se mencionó anteriormente, existen algunos ejercicios realizados por el DANE que incluyen este sector, por ejemplo, el primer censo sobre recursos informáticos realizado por el DANE en 1983. Posteriormente, en 2001 se realizó la primera encuesta de acceso y uso de Tecnologías de la Información y las Comunicaciones en Hogares y empresas (sectores Industria, Comercio, Servicios y Educación).

Ahora bien, desde 2007, el DANE calcula para Colombia los indicadores básicos de tenencia y uso de las TIC en hogares y personas a través de unos módulos insertados en la Gran Encuesta Integrada de Hogares (GEIH) (hasta 2011) y en la Encuesta de Calidad de Vida (ECV) (desde 2012).

Teniendo en cuenta las necesidades de medir las TIC y la transformación digital, el Gobierno Nacional ha creado la Política Nacional para la Transformación Digital e Inteligencia Artificial planteada en el documento CONPES 3975 de 2019, con la cual busca disminuir las barreras que impiden el desarrollo digital en las empresas, así como también, señala que el MINTIC deberá aplicar a nivel nacional la Encuesta TIC en conjunto con el DANE para medir el avance de la transformación digital en el país, con la desagregación departamental para hogares y nacional para empresas, siguiendo el marco de trabajo definido por la OCDE.

Desde mediados de 2018 y durante 2019 se desarrollaron mesas de trabajo con personal técnico del Ministerio de Tecnologías de la Información y Comunicaciones (MINTIC), con el objetivo de diseñar una gran encuesta que no sólo midiera los indicadores básicos sobre acceso y uso de TIC que actualmente publica el DANE, sino que permitiera analizar y atender nuevos requerimientos internacionales sobre medición de la sociedad de la información y generar indicadores especializados que permitan promover el proceso de transformación digital en Colombia. Así mismo, el documento CONPES 3975 de 2019, en su línea de acción 3, indica que el MINTIC y el DANE deberán aplicar la encuesta a nivel nacional con un nivel de desagregación departamental para hogares.

Fruto de esas mesas de trabajo se suscribieron los convenios interadministrativos 781 de 2019 y 899 de 2021 suscritos entre el DANE y el MINTIC, lo cual permitió la realización de dos encuestas especializadas sobre tenencia y uso de TIC en hogares ENTIC Hogares para los periodos de referencia 2020 y 2021.

2.1.5. Definición de variables y construcción de indicadores estadísticos

Las principales variables de la ENTIC Hogares son las siguientes:

- **Tenencia de bienes y servicios TIC en los hogares:** posesión de TIC para el periodo de referencia y características de la conexión.
- **Uso de bienes y servicios TIC en los hogares:** utilización de modo habitual de TIC. Se mide en términos del principal uso y la frecuencia de uso de TIC.
- **Servicios del hogar:** conexión a servicios públicos, tales como energía, telefonía e internet, tanto fijo como móvil.
- **Seguridad y uso responsable de internet:** hábitos y prácticas preventivas y correctivas de las personas para proteger sus conexiones a servicios de telefonía e internet tanto fijo como móvil, redes sociales, datos personales, entre otros.

Por su parte, los indicadores de tenencia y uso de TIC preservan un marco teórico fundamental que corresponde a los principales acuerdos alcanzados por la comunidad de personas expertas, nacionales e internacionales sobre diseño, aplicación e interpretación de encuestas nacionales de TIC, por lo que sigue los lineamientos del Manual para la medición de acceso y uso de TIC en hogares y personas de la ITU (2020). A continuación, se presentan los indicadores principales:

- Indicadores de tenencia de TIC a nivel de hogar.
- Indicadores de uso de TIC a nivel de persona.
- Indicadores de seguridad en TIC.

Indicadores de tenencia de TIC a nivel de hogar

Hogares con servicio de teléfono fijo (HSTF): corresponde al porcentaje de hogares que cuentan con el servicio de teléfono fijo, corriente o tradicional.

$$HSTF = \frac{\text{Hogares que cuentan con servicio de teléfono fijo}}{\text{Total de hogares}} * 100$$

Hogares con servicio de teléfono celular (HSTC): corresponde al porcentaje de hogares en los que alguno de sus miembros cuenta con el servicio de teléfono celular.

Hogares que poseen computador de escritorio, portátil o tableta (HPC): corresponde al porcentaje de hogares que cuentan con computador de escritorio, portátil o tableta.

$$HPC = \frac{\text{Hogares que poseen un computador de escritorio, portátil o tableta}}{\text{Total de hogares}} * 100$$

Hogares que poseen televisor convencional a color, LCD, plasma o LED (HPTV): corresponde al porcentaje de hogares que cuentan con televisor convencional a color, LCD, plasma o LED.

$$HPTV = \frac{\text{Hogares que poseen televisor convencional a color, LCD, plasma o LED}}{\text{Total de hogares}} * 100$$

Hogares que poseen conexión a Internet (HPCI): corresponde al porcentaje de hogares que poseen conexión a internet.

$$HPCI = \frac{\text{Hogares que poseen conexión a Internet}}{\text{Total de hogares}} * 100$$

Personas que poseen teléfono celular (PPTC): corresponde al porcentaje de personas de 5 años y más que poseen teléfono celular.

$$PPTC = \frac{\text{Personas de 5 años y más que poseen telefono celular}}{\text{Total de Personas de 5 años y más}} * 100$$

Indicadores de uso de TIC a nivel de persona

Personas que usan teléfono celular (UTC): corresponde al porcentaje de personas de 5 años y más que usan teléfono celular.

$$UTC = \frac{\text{Personas de 5 años y más que usan telefono celular}}{\text{Total de Personas de 5 años y más}} * 100$$

Personas que usan computador en cualquier lugar (UCCL): corresponde al porcentaje de personas de 5 años y más que manifiestan usar computador (de escritorio, portátil o tableta)

$$UCCL = \frac{\text{Personas de 5 años y más que usan computador (de escritorio portátil o tableta)}}{\text{Total de Personas de 5 años y más}} * 100$$

Personas que escucharon la señal de radio dentro del hogar (USRH): corresponde al porcentaje de personas de 5 años y más que escucharon la señal de radio dentro del hogar.

$$USRH = \frac{\text{Personas de 5 años y más que escuchan la señal de radio dentro del hogar}}{\text{Total de Personas de 5 años y más}} * 100$$

Personas que usan Internet en cualquier lugar (UICL): corresponde al porcentaje de personas de 5 años y más que manifiestan usar Internet en cualquier lugar y desde cualquier dispositivo.

$$UICL = \frac{\text{Personas de 5 años y más que usan Internet en cualquier lugar}}{\text{Total de Personas de 5 años y más}} * 100$$

Personas que usan herramientas de Inteligencia Artificial (UIA): corresponde al porcentaje de personas de 5 años y más que manifiestan Utilizan herramientas de Inteligencia artificial.

$$UIA = \frac{\text{Personas de 5 años y más que usan herramientas de Inteligencia Artificial}}{\text{Total de Personas de 5 años y más}} * 100$$

Hogares que usan el servicio de televisión por cable, satelital o IPTV (HUTV): Corresponde al porcentaje de hogares que usan el servicio de televisión por cable, satelital o IPTV.

$$HUTV = \frac{\text{Hogares que usan el servicio de televisión por cable, satelital o IPTV}}{\text{Total de hogares}} * 100$$

Personas que usan redes sociales (URS): Corresponde al porcentaje de personas de 5 años y más que usan redes sociales.

$$URS = \frac{\text{Personas de 5 años y más que usan redes sociales}}{\text{Total de Personas de 5 años y más}} * 100$$

Personas que efectuaron compra de bienes y/o servicios en línea (UCBS): Corresponde al porcentaje de Personas de 5 años y más que compraron bienes y/o servicios en línea.

$$UCBS = \frac{\text{Personas de 5 años y más que compraron bienes y/o servicios en línea}}{\text{Total de Personas de 5 años y más}} * 100$$

Indicadores de seguridad TIC

Personas que usan dispositivos protegidos para acceder a Internet (UPI): Corresponde al porcentaje de personas de 5 años y más que protegen sus dispositivos de amenazas del internet.

$$UPI = \frac{\text{Personas de 5 años y más que cuentan con un dispositivo protegido para acceder a internet}}{\text{Total de Personas de 5 años y más}} * 100$$


Frecuencia de incidentes con el uso las TIC (ITIC): corresponde al porcentaje de personas de 5 años y más que han tenido problemas durante el uso de TIC.

$$\% \text{ relevancia de los incidentes por tipo} = \frac{\text{Total de incidentes por tipo}}{\text{Total de problemas reportados}} * 100$$

2.1.6. Resultados estadísticos

Los resultados de la ENTIC Hogares se publican a través de un informe técnico y anexos estadísticos en formato Excel con los principales cuadros de salida. Además, se publican los microdatos anonimizados a través de la Sala de Procesamiento Especializado Externo (SPEE).

Los cuadros de salida generados en la operación estadística incluyen las variables representativas de cada módulo de la encuesta a nivel de hogares y personas según temática. En el Anexo A se incluye la relación de cuadros de salida que serán publicados en los anexos estadísticos.

 DANE	METODOLOGÍA GENERAL	CÓDIGO: DSO-ENTIC_H-MET-001 VERSIÓN: 1
PROCESO: Producción Estadística	OPERACIÓN ESTADÍSTICA: Encuesta de Tecnologías de la Información y las Comunicaciones en Hogares- ENTIC_H	

Cuadros de salida

Figura 2. Esquema de cuadros de salida ENTIC Hogares



Fuente: Departamento Administrativo Nacional de Estadística.

El proceso inicia con el diseño de los cuadros y sus especificaciones, las cuales son elaboradas por el grupo temático y posteriormente enviadas al área de sistemas para su programación y procesamiento. Luego de este paso son enviadas de vuelta al área temática para su revisión y verificación.

Para la ENTIC Hogares, los cuadros de salida se presentan con la siguiente desagregación: total nacional, cabecera y centros poblados-rural disperso y total departamental. Los resultados presentados comprenden tanto valores absolutos como sus respectivos valores porcentuales. En la Figura 2 se relaciona un esquema con la relación de cuadros de salida que serán publicados en los anexos estadísticos (Para más detalle de esta lista de Cuadros de salida ver Anexo A).

2.1.7. Estándares estadísticos utilizados

La ENTIC hace uso de las siguientes nomenclaturas y clasificaciones:

División Político-Administrativa de Colombia (DIVIPOLA): Estándar nacional que codifica y lista las entidades territoriales a saber: departamentos, municipios, corregimientos departamentales, así como los centros poblados, tanto inspecciones de policía, como caseríos y corregimientos municipales en el área rural. El objeto primordial de contar con la totalidad de estas unidades inventariadas es identificarlas mediante un código numérico y asociarles información temática de diferentes fuentes, mediante el uso de sus identificadores.

Data Documentation Initiative DDI y Dublin Core: Mediante el uso del software Nesstar Publisher, la ENTIC Hogares utiliza dichos estándares para documentar la operación, publicar los

micro y metadatos y vincular los documentos de soporte (metodología, ficha metodológica y manuales) para consulta y descarga de los usuarios en el portal Archivo Nacional de Datos (ANDA).

Para el proceso de elaboración de cada uno de los documentos de la ENTIC Hogares, se tienen como base cada una de las **Guías vigentes y recomendadas por el Sistema Estadístico Nacional (SEN)**.

2.1.8. Diseño del cuestionario

En coherencia con el objetivo de obtener la información necesaria para la actualización de los indicadores de tenencia y uso de las TIC a nivel de viviendas, hogares y personas, cuyos resultados sirvan como insumo para la definición de políticas que permitan diseñar y ejecutar planes sociales.

El formulario diseñado para la ENTIC Hogares incluye cinco (5) módulos permanentes de la operación estadística.

Módulo I: Localización y vivienda.

En este módulo se hace una identificación de la ubicación de las viviendas y los hogares, así como un control de calidad sobre la recolección de las encuestas y preguntas a nivel de vivienda.

Módulo II: Tenencia de Bienes y Servicios TIC en el Hogar.

En este módulo se capta información relacionada con la tenencia Bienes y Servicios TIC en el hogar (Tenencia telefónica, computadores, televisores y dispositivos tecnológicos; servicios de conexión a internet, televisión y radio, máxima velocidad de internet).

Módulo III: Características y composición del hogar (para todas las personas del hogar)

En este módulo se busca identificar las personas que conforman el hogar y establecer su parentesco con el jefe(a) del hogar. Además, se realiza una caracterización demográfica de todos los integrantes en aspectos como sexo, edad, estado civil, autorreconocimiento étnico, lugar de nacimiento y aspectos de migración.

Módulo IV: Uso de TIC (para todas las personas de 5 años y más)

El objetivo del módulo es medir el uso de las Tecnologías de la Información y las Comunicaciones (TIC) por parte de las personas, con énfasis en el uso del computador (de escritorio, portátil y tableta), la internet y la telefonía móvil celular. Entre otros aspectos, se indaga por lugares de acceso, frecuencia de uso, actividades realizadas a través de Internet, utilización de herramientas de Inteligencia Artificial (IA), tenencia y uso de la telefonía móvil celular y uso de la señal de radio.

Módulo V: Seguridad y uso responsable del Internet (para todas las personas de 5 años y más)

En este módulo se busca identificar los mecanismos que usan las personas durante el uso del Internet para proteger de posibles problemas como captura de virus, pérdidas económicas y divulgación no consentida de videos o imágenes, entre otras. Así mismo, se indaga por los incidentes de seguridad digitales ocurridos y las acciones que toman las personas después de haber sido víctima de uno de estos.

Así mismo, con el fin de garantizar la coherencia de la información recolectada y la secuencia lógica de las preguntas, el cuestionario fue diseñado en conjunto con las especificaciones de validación donde se documentan las condiciones para mantener la consistencia entre las variables y por lo tanto, el correcto flujo de la información, junto con la caracterización de estas.

2.1.9. Normas, especificaciones o reglas de edición e imputación de datos

El proceso de generación de normas de validación y consistencia para las preguntas de la ENTIC Hogares comprende en la definición de flujos, rangos y valores válidos de cada variable del formulario por parte del grupo temático. El propósito de su generación es asegurar que, por ejemplo, las preguntas del formulario electrónico se diligencien de acuerdo con los rangos de edad correspondientes para cada capítulo, se apliquen los respectivos flujos o saltos de preguntas y se valide la consistencia de un capítulo a otro y entre las mismas preguntas.

Las normas definidas se consolidan en un formato que el grupo temático entrega al equipo de sistemas para su inclusión en el programa de captura y control de inconsistencias de la información.

Las normas no incluyen consideraciones para la imputación de datos, por no ser una actividad que se adelanta en esta operación estadística.

2.2. DISEÑO ESTADÍSTICO

El diseño estadístico reúne varios aspectos entre los cuales se destacan la formulación del marco estadístico, el universo, la población, las unidades estadísticas, los periodos y el diseño muestral, de modo que, para la aplicación en 2024, se actualizó el marco muestral con el censo del 2018; adicionalmente, se hicieron ajustes en algunos aspectos metodológicos.

2.2.1. Universo de estudio

El universo para la Encuesta Tecnologías de la Información y Comunicación en Hogares (ENTIC) Hogares está conformado por la población civil no institucional residente en todos los municipios del territorio nacional, excluyendo la parte rural de San Andrés.

2.2.2. Población objetivo

La población objetivo de la operación estadística corresponde a los hogares conformados por personas residentes en el territorio nacional; sin embargo, se excluye aquella población para la cual el desarrollo de la encuesta genera sobrecostos y que al mismo tiempo su eliminación no afecta los resultados generales de la operación estadística. De esta manera la operación estadística excluye la población de Providencia y el centro poblado y rural disperso de San Andrés.

2.2.3. Cobertura geográfica

La ENTIC Hogares, cuenta con un cubrimiento nacional y departamental, para el caso de individuos y hogares. Por el diseño muestral no es posible tener desagregación de información a nivel municipal.

Antes de describir la cobertura y desagregación geográfica de la ENTIC Hogares es necesario definir los conceptos de cabeceras, centros poblados y rural disperso.

Cabecera: *Delimitación geográfica definida por el DANE para fines estadísticos, alusiva al área geográfica delimitada por el perímetro censal. A su interior se localiza la sede administrativa del municipio, es decir la alcaldía.* Departamento Administrativo Nacional de Estadística - DANE. (2018). Censo Nacional de Población y Vivienda CNPV, Manual de Conceptos 2018.

Centro poblado: *Concepto construido por el DANE con fines estadísticos, para la identificación y localización geográfica de núcleos o asentamientos de población. Se define como una concentración mínima de veinte viviendas contiguas, vecinas o adosadas entre sí, ubicada en el área resto municipal o en un área no municipalizada (corregimiento departamental). Contempla los núcleos de población de los corregimientos municipales, inspecciones de policía y caseríos. Dicha concentración presenta características tales como la delimitación de vías vehiculares y peatonales.* Departamento Administrativo Nacional de Estadística - DANE. (2018). Censo Nacional de Población y Vivienda CNPV, Manual de Conceptos 2018.

Este concepto considera:

Caserío: *Definición construida por el DANE para fines estadísticos, que corresponde a un centro poblado ubicado generalmente al lado de una vía principal y que no tiene autoridad civil. El límite censal está definido por las mismas viviendas que constituyen el conglomerado.* Departamento Administrativo Nacional de Estadística (DANE). Manual Técnico Censo general 2005

Inspección de Policía: El DANE, para fines estadísticos, la define como conglomerado de viviendas ubicadas en el área resto municipal que tiene por autoridad principal al inspector de policía. Instituto Geográfico Agustín Codazzi (IGAC). Glosario Diccionario geográfico de Colombia.

Corregimiento municipal: *Tipo de centro poblado, ubicado en el área rural de un municipio, el cual incluye un núcleo de población y está considerado en los Planes de Ordenamiento Territorial (POT).* Adaptado de: Ley 136 de 1994. Por la cual se dictan normas tendientes a modernizar la organización y el funcionamiento de los municipios. Diario Oficial No. 41.377 de 2 de junio de 1994. y Ley 1551 de 2012. Por la cual se dictan normas para modernizar la organización y el funcionamiento de los municipios. Diario Oficial No. 48.483 de 6 de julio de 2012.

Rural disperso: *Delimitación geográfica definida por el DANE para fines estadísticos, comprendida entre el perímetro censal de las cabeceras municipales y de los centros poblados, y el límite municipal. Se caracteriza por la disposición dispersa de viviendas y de explotaciones agropecuarias existentes en ella. Corresponde al territorio que no forma parte ni de la cabecera municipal (clase 1) ni de los centros poblados (clase 2). Se caracteriza por objetos y elementos relacionados con la agricultura, predios de descanso o recreo, usos mineros o extractivos. El número de unidades residenciales por área es menor a las zonas urbanas.* Departamento Administrativo Nacional de Estadística - DANE. (2018). Censo Nacional de Población y Vivienda CNPV, Manual de Conceptos 2018.

Resto: *Corresponde al área geográfica cubierta por el centro poblado y rural disperso.* (Departamento Administrativo Nacional de Estadística).

2.2.4. Desagregación geográfica

La ENTIC Hogares tendrá una desagregación geográfica nacional por cabecera y centros poblados - rural disperso y Total por departamentos.

Desagregación Geográfica	Descripción
Total Nacional	Total nacional y desagregado por cabecera y centro poblado – rural disperso.
Departamental	Total por departamento

2.2.5. Desagregación temática

La desagregación temática se divide en dos grupos principales:

1. Indicadores de tenencia de TIC a nivel en hogares.
2. Indicadores de uso de TIC por parte de personas de 5 y más años.

El tamaño de muestra de la ENTIC Hogares 2024 permite presentar resultados desagregados por sexo y otros grupos poblacionales de interés, dependiendo de los errores de muestreo asociados a cada desagregación, por este motivo no se recomienda usar desagregaciones cuyos indicadores posean coeficientes de variación superiores al 15%.

2.2.6. Fuentes de datos

La ENTIC Hogares obtiene la información de la fuente primaria, es decir directamente de los hogares, a través del jefe del hogar y de las personas con encuesta directa.

El marco muestral lo constituye el listado de viviendas, hogares y personas y el inventario cartográfico obtenido de la información del Censo Nacional de Población y Vivienda (CNPV) de 2018, así como las actualizaciones del período intercensal.

2.2.7. Unidades estadísticas

El conjunto de unidades empleadas para el diseño, la medición y la presentación de resultados son:

- a) **Unidades de observación:** corresponde a las viviendas, los hogares y las personas de 5 o más años.

- b) **Unidades de análisis:** es representada por las viviendas, los hogares y personas de 5 o más años.
- c) **Unidades de muestreo:** para esta operación estadística se considera un muestreo en 2 etapas, donde, las unidades primarias de muestreo se relacionan con la primera etapa y las unidades secundarias con la segunda etapa. De tal manera que las unidades de muestreo son las siguientes:
- **Unidades Primarias de Muestreo (UPM):** para este estudio las unidades primarias de muestreo corresponden a los municipios del país.
 - **Unidades Secundarias de Muestreo (USM):** hace referencia a los segmentos o conglomerados de aproximadamente 10 viviendas contiguas, también denominados medidas de tamaño (MT), ubicadas tanto en las cabeceras como en el resto de cada municipio.

Nota: En las USM o MT se mide todas las viviendas, hogares y personas, por lo que no corresponden a unidades muestrales.

2.2.8. Período de referencia

La ENTIC Hogares recoge información de diferentes temas, razón por la cual, dentro de cada módulo existen variables que se estudian con frecuencias diferentes, teniendo en cuenta que existen hechos que por su magnitud e importancia presentan mayor recordación que otros.

Los períodos de referencia que se manejan en la ENTIC Hogares:

- **Últimos 12 meses:** contados a partir del periodo de recolección.
- **Semana pasada o de referencia:** semana calendario (lunes a domingo) inmediatamente anterior a la semana en la que se efectúa la encuesta. Es utilizado para obtener información sobre la fuerza de trabajo.

2.2.9. Periodo de recolección y frecuencia

La recolección toma en promedio 3 meses a partir del inicio del operativo y su frecuencia es anual.

2.2.10. Marco estadístico muestral

La muestra de la ENTIC hogares de 2024 se selecciona en dos etapas de muestreo. Cada etapa tiene un marco muestral asociado, la primera se usa un marco de municipios con una estructura de estratificación ajustada a las necesidades de las investigaciones sociales y la segunda corresponde con medidas de tamaño o segmentos estratificada por dominio geográfico y tiene su origen en el marco geoestadístico nacional, producto del Censo Nacional de Población y Vivienda (CNPV) del 2018.

Marco Muestral de municipios

El marco de municipios es un marco que enlista cada municipio del país (vigencia censo 2018), de modo que cuenta con información auxiliar de las poblaciones de cabecera y centro poblado – rural disperso, indicaciones del índice de pobreza multidimensional, porcentaje de centro poblado y rural disperso. De igual forma, incluye los estratos de muestreo de la primera etapa.

Marco Muestral de Medidas de Tamaño o Segmentos

Es un marco construido a partir de la información estadística de densidad de viviendas en las manzanas y secciones rurales del marco geoestadístico nacional, sirve para identificar y ubicar geográficamente las USM, por lo tanto, enlista cada conglomerado de 10 viviendas del país, asociándole la manzana o sección rural a la que pertenece.

Para construir el marco se toma el número de viviendas que hay en cada manzana o sección rural del marco geoestadístico nacional y luego se divide en tamaños de aproximadamente 10 viviendas. Como resultado se obtiene el número de medidas de tamaño teóricas asociadas a cada manzana y sección rural del país. Con esta información se crea un marco nuevo de muestreo donde las filas son los conglomerados; por ejemplo, si una manzana tiene 30 viviendas entonces al dividir en 10, significa que se tienen 3 medidas de tamaño y que, en consecuencia, en el marco de conglomerados, van a aparecer 3 filas para representar los 3 conglomerados.

Con este marco se selecciona directamente la muestra de segmentos en la segunda etapa de esta encuesta, luego de la selección se identifican las manzanas o secciones asociadas a la muestra y se procede a crear un marco de viviendas segmentado. Para el caso de cabeceras y centro poblado se desarrolla un operativo de campo dirigido a las manzanas asociadas, a estas medidas de tamaño, donde se procede a enlistar todas las viviendas contenidas en las manzanas. Este marco se segmenta en medidas de tamaño de 10 viviendas. En lo que respecta a la zona rural dispersa, cada sección se segmenta creando polígonos que, según información de foto satelital contengan un

aproximado de 10 viviendas. En ambos casos la segmentación se enumera y posteriormente se encuesta la medida de tamaño elegida en la primera parte.

2.2.11. Diseño muestral

La ENTIC-Hogares es una encuesta dirigida a hogares y personas, por lo tanto, se establece una estrategia muestral adecuada para acceder a mediciones de dichas unidades de análisis con procedimientos que aseguran estimadores aproximadamente insesgados de los parámetros bajo una precisión y confianza preestablecida de antemano. A continuación, se describirá en detalle el plan muestral.

Tipo de muestreo:

Teniendo en cuenta los objetivos y las características de la encuesta se utiliza una muestra probabilística, multietápica, estratificada y de conglomerados.

- a) Probabilística: Se satisfacen los siguientes criterios, a saber:
1. Se conoce de ante mano todas las posibles muestras.
 2. Cada posible muestra tiene su respectiva probabilidad de selección.
 3. El procedimiento da a cada elemento del universo de estudio una probabilidad de inclusión conocida y mayor que cero.
 4. Se tiene un mecanismo aleatorio dentro del cual a cada posible muestra tiene una probabilidad no nula de selección.
- b) Multietápica: Para lograr la muestra de viviendas, hogares y personas se seleccionaron secuencialmente las unidades de muestreo (UPM y USM) en dos etapas
- En la primera se seleccionan municipios con un muestreo sin reemplazamiento proporcional al tamaño.
 - En la segunda se seleccionan conglomerados de 10 viviendas con un muestreo sistemático con un ordenamiento geográfico previo.
- c) Estratificada: Para la ENTIC Hogares se utiliza el sistema de estratificación de municipios creado desde 2020 para UPM y otra basada en un criterio geográfico para USM.

Respecto a la estratificación de UPM, en la primera etapa de la muestra se generan grupos de municipios dentro de cada departamento. En el primer grupo de cada departamento se establece el

estrato de inclusión forzosa que contiene un único municipio que corresponde a la capital de departamento.

El segundo grupo lo componen los municipios considerados como áreas metropolitanas de las capitales de departamento. Las otras agrupaciones de municipios se hicieron a partir de un Análisis de Componentes Principales (ACP) y un posterior agrupamiento tomando en cuenta la proyección de los municipios sobre los nuevos ejes factoriales y la población. Para lograrlo, se tomaron los resultados del Censo 2018, el índice de pobreza multidimensional, sus 15 indicadores y el porcentaje de urbanización. Con el propósito de buscar un diseño muestral auto ponderado, los grupos consolidados tienen aproximadamente la misma población, aunque esta característica no se puede asegurar; por tal motivo, van a existir unos estratos más grandes que otros y algunos de inclusión forzosa de municipios con una gran población, es decir, que no corresponden al tamaño promedio de los estratos.

En la segunda etapa la estratificación se hace a nivel de cabecera, centro poblado y rural disperso. Por último, se hace una estratificación implícita de los segmentos al ordenar el marco según sector, sección y manzana y luego seleccionar a través de un muestreo sistemático la muestra de USM. Este procedimiento garantiza que la muestra quede uniformemente distribuida sobre las áreas geográficas.

d) De conglomerados: Los conglomerados están definidos por la unidad secundaria de muestreo (USM) y corresponden a 10 viviendas (contiguas) en promedio donde se encuestan todas las viviendas, los hogares y las personas que lo conforman.

Cálculo del tamaño de muestra

Para establecer un tamaño de muestra, en primera instancia, se debe identificar un indicador relevante en la investigación al cual se le asigne un tamaño de muestra suficiente para que el estimador de este indicador cumpla con ciertas características de precisión y confianza.

En los casos donde el diseño muestral es complejo, además de lo anterior, se debe tomar como referencia un muestreo aleatorio simple de elementos (MAS) estableciendo un tamaño de muestra primario, el cual deberá ser ajustado por un efecto del diseño muestral (Deff) y otros parámetros como la tasa de no respuesta esperada y las subdivisiones de la población donde, en esta última, se procura garantizar suficiente muestra para su respectiva desagregación. Se invita al lector a revisar el desarrollo teórico del tema en (Cochran, 1977) o en las metodologías publicadas por Naciones Unidas sobre encuestas de hogares (Naciones Unidas, 2008).

En lo que respecta a esta encuesta se busca establecer tamaños de muestra para indicadores de proporción, con lo cual se define el tamaño de muestra como:

$$n = \frac{N(e^2 + PQz_{1-\alpha}^2)}{Ne^2 + PQz_{1-\alpha}^2} * deff * T$$

Donde,

N : Proyección de personas para la población objetivo

e : Margen de error ajustado para un Cve del 5%

P : Prevalencia referente para la encuesta, 10% en alineación con otras encuestas sociales

$z_{1-\alpha}$: Percentil de una normal estándar para una confianza del 95%, valor 1.96

$deff$: Efecto del diseño por selección de conglomerados, asumido como 2.05 para MT de 10 viviendas

T : Ajuste por no respuesta, si t es el porcentaje de no respuesta se calcula el factor de ajuste como.

$$T = 1/(1 - t)$$

De acuerdo a esto se calculan los valores de tamaño de muestra por cada ciudad, y teniendo en cuenta que resulta conveniente tener una muestra balanceada ya que permite estimaciones de otros indicadores por ciudad mejor distribuidos se logra la siguiente distribución por segmentos de la muestra ENTIC de 2024.

En la construcción de número de segmentos a seleccionar se divide a n por el número de personas por hogar (PPH), por el número de hogares por vivienda (PHV) y en 10.

Es decir:

$$\text{Número de segmentos a encuestar} = n/(PPH * PHV * 10)$$

De acuerdo a esto se estiman los siguientes tamaños de muestra en número de segmentos por departamento, así:

Departamento	Municipios (UPM)	Segmentos (USM)		
		Cabeceras	Centros Poblados	Rural Disperso
05-Antioquia	41	211	50	95
08-Atlántico	11	117	18	27
11-Bogotá, D.C.	1	150	3	15
13-Bolívar	14	105	29	26
15-Boyacá	26	273	2	72
17-Caldas	15	113	12	32
18-Caquetá	10	173	12	53
19-Cauca	15	69	29	70
20-Cesar	12	109	24	26
23-Córdoba	14	77	23	42
25-Cundinamarca	23	132	17	46
27-Chocó	13	60	29	51
41-Huila	13	92	14	43
44-La Guajira	10	76	23	54
47-Magdalena	11	89	23	15
50-Meta	11	124	9	31
52-Nariño	16	73	24	52
54-Norte de Santander	14	121	10	23
63-Quindío	9	134	10	23
66-Risaralda	10	108	8	29
68-Santander	17	133	13	55
70-Sucre	13	132	47	55
73-Tolima	16	112	11	43
76-Valle del Cauca	18	115	27	33
81-Arauca	7	64	3	27
85-Casanare	14	95	4	34
86-Putumayo	12	64	17	54
88-Archipiélago de San Andrés	1	111	0	0
91-Amazonas	11	44	18	31
94-Guainía	8	37	7	26
95-Guaviare	4	65	8	30
97-Vaupés	6	33	4	35
99-Vichada	4	27	9	48

Fuente: Departamento Administrativo Nacional de Estadística

En total la ENTIC-Hogares 2024 tendrá presencia en 420 municipios del país, en los cuales se trabajarán un total de 5271 segmentos y se espera encuestar cerca de 58.000 hogares.

Selección de muestra

En la primera etapa, la UPM o municipio se selecciona dentro de cada estrato de muestreo con un algoritmo de selección proporcional al tamaño (PPT) con covariable de tamaño dada por el censo total de personas para 2018.

Para la segunda, se ordena el marco por sector, sección y manzana luego, se hace una segmentación del marco muestral, de tal manera que la USM tenga aproximadamente 10 viviendas en promedio. Luego dentro de cada estrato de la etapa se hace una selección sistemática de segmentos con único arranque. Para asegurar un tamaño de muestra fijo se usa el método de "Fractional Interval Method".

Todos los procedimientos se hacen con PROC SURVEYSELECT del software SAS (SAS Institute Inc, 2019). Por lo cual, las especificaciones de algoritmos usados se pueden revisar en la página web del programa.

Mantenimiento de muestra y manejo de novedades muestrales:

En el desarrollo normal de la operación estadística, se identifican diversos aspectos que pueden causar sesgos tales como la desactualización de marco, la imposibilidad de acceder a la muestra seleccionada con el procedimiento probabilístico predefinido debido a inconvenientes en campo y al agotamiento de la fuente.

La desactualización de marco se genera con el tiempo debido a que en las áreas geográficas la población puede variar. Para controlar la desactualización del marco se procede a tomar en cuenta estos cambios con un factor de corrección. En la medida que las manzanas varíen en su número de viviendas las medidas de tamaño también lo harán en proporción ajustándolas en su factor de expansión con un factor de corrección que toma en cuenta el cambio del número de viviendas de la manzana respecto a lo medido en el censo. La fórmula se define como el número de segmentos actuales sobre el número de segmentos que aparecían en el marco al momento de seleccionar la medida de tamaño. Posteriormente, se realiza un proceso en el cual, con la información del recuento se va actualizando el marco de medidas de tamaño en lo que respecta al número de segmentos y luego, se ajusta la muestra según el resultado de la actualización.

Con respecto a la de pérdida de muestra, lo que se solicita es hacer revisitas en la medida de lo posible. En los casos donde se presentan novedades como pérdida de segmentos por problemas de

orden público o por rechazo de segmentos, se tiene protocolos que se siguen para evitar la pérdida; sin embargo, en los casos donde es inevitable la pérdida de segmentos se opta por reemplazarlos con segmentos lo más parecidos posible al segmento original (estos casos son escasos).

2.2.12. Ajustes de cobertura (o ajuste de cobertura por no respuesta)

El ajuste de cobertura se realiza en los casos donde se pierden viviendas, hogares y personas, el procedimiento describe la construcción de un factor de cobertura (FCOB). Este factor restituye a nivel segmento las pérdidas ocasionadas por el operativo de campo. Es importante resaltar que, para evitar sesgos de selección, cuando se pierde de forma parcial o total la información de una persona dentro del hogar, el hogar se da como perdido y se toma en cuenta en el factor de cobertura.

Ponderadores y metodología de estimación de parámetros y varianzas de estimadores:

Los ponderadores usados en las encuestas de hogares que hace el DANE son construidos por etapas de la siguiente manera:

Factor de corrección de desactualización de marco (PSUB):

Es un factor a nivel de segmento cuyo objetivo es nivelar el peso con el que cuenta el segmento al que debería tener por el aumento o disminución de viviendas en el momento actual. En este sentido, se puede ver como una etapa de la muestra, en donde el conglomerado final es partido en varios conglomerados de los cuales se selecciona sólo uno. En este sentido, se expande entonces con el factor de sub-muestreo para de nuevo restituir el número de viviendas del conglomerado final.

Como ejemplo supóngase que se tiene una manzana de 20 viviendas en el censo 2018 (2 medidas de tamaño). Para el 2020 el número de viviendas encontradas es 50 (5 medidas de tamaño). Es claro que, al segmentar estas 50 viviendas en 2, cada conglomerado final tendría un tamaño de 25 viviendas; sin embargo, encuestar 25 viviendas es costoso, lo ideal es mantenerse en un tamaño de muestra fijo, así que se prefiere encuestar 10 viviendas y expandirlas a 25 con el peso de sub-muestreo. Es fácil verificar que el factor es aproximadamente $25/10$ (en termino de viviendas) $=5/2$ (en termino de medidas de tamaño) $=2,5$.

De esta forma, el peso de sub-muestreo se construye contrastando el número de medidas de tamaño que se definieron en la manzana a partir de la información de viviendas del censo 2018 (MTs teóricas), con las encontradas en campo en el momento de hacer la encuesta. Su fórmula es la siguiente:

$$PSUB = \frac{\text{Número de MTs actuales en la manzana}}{\text{Número de MTs teóricas en la manzana}}$$

Donde el número de MTs teóricas de la manzana es el entero más cercano de la razón entre el número de viviendas del censo 2018 y 10. Mientras el número de MTs actuales en la manzana es el entero más cercano del número de viviendas encontradas en el recuento y 10.

En el proceso de uso de este factor cabe aclarar que no se aplica de manera generalizada, sino que se restringe a unos límites con un mínimo de 0.8 y un máximo de 1.5, para evitar que el factor genere desproporciones estructurales en la medición de los indicadores.

Factor de cobertura (FCOB):

Es un factor a nivel de segmento que corrige la pérdida de viviendas, hogares y personas dentro del segmento. Su fórmula se establece de la siguiente forma:

- Controlando el tamaño del segmento: En la práctica los segmentos aparecen de diferentes tamaños, aunque se definen segmentos de 10 viviendas desde el diseño muestral. Para controlar este fenómeno se nivela la representación del segmento a 10 viviendas por segmento.

$$fcob1 = \frac{10}{\text{Número de viviendas encuestadas efectivas en el segmento}}$$

- Se ajusta las viviendas vacantes: La naturaleza de las áreas geográficas debe ser representada. El número de viviendas vacantes debe ser descontado del número de viviendas del segmento, para lograrlo se usa la siguiente relación.

$$fcob2 = \frac{\text{Número de viviendas en el segmento} - \text{viviendas vacantes}}{\text{Número de viviendas en el segmento}}$$

El factor de cobertura final es la multiplicación de los dos anteriores.

$$FCOB = fcob1 \cdot fcob2$$

Es fundamental resaltar que el número de viviendas en el segmento representa al número de viviendas identificadas en el recuento y que no se excluyen ni las vacantes ni las perdidas en el

operativo. Así mismo, el número de viviendas encuestadas efectivas en el segmento se refiere a las viviendas con encuesta completa que terminan en la base de datos.

El factor también posee una restricción de un máximo de 1.5 para evitar los sesgos generados por sobre-expansión de algunas unidades Muestrales.

Factor de segmento a municipio (FMUNI):

Es un factor asociado a todos los segmentos de un municipio en su cabecera o en su resto (siempre la operación se hace de forma independiente para ambas clases). El factor expande los individuos de la muestra al total de individuos del municipio en su cabecera o resto y tiene la siguiente formula:

$$FMUNI = \frac{\text{Número de personas en el mpio (cabecera o resto)}}{\text{Número de personas en la muestra del mpio (cabecera o resto)}}$$

Factor de municipio a estrato (FESTR):

Es un factor asociado a todos los municipios de un estrato de muestreo en la primera etapa de la muestra (Estratificación de UPMs). La tarea del factor consiste en expandir la población de los municipios seleccionados en la primera etapa a la población de todos los municipios dentro de sus respectivos estratos. Como el anterior factor, este ponderador se aplica independientemente según cabecera y resto.

$$FESTR = \frac{\text{Número de individuos de los municipios pertenecientes al estrato según cabecera y resto}}{\text{Número de individuos que conforman la muestra en el estrato según cabecera resto}}$$

Factor básico (d_k):

Es el factor resultado de la multiplicación de los anteriores factores de expansión. Se pone el subíndice k para identificar la vivienda k y seguir la notación de los artículos de calibración (Deville & Särndal, 1996).

$$d_k = PSUB_K \cdot FCOB_K \cdot FMUNI \cdot FESTR$$

Nótese que hasta este punto las viviendas, hogares y personas dentro de un mismo segmento comparten el mismo factor de expansión básico.

Factor calibrado (w_k):

Los factores calibrados son el resultado de un proceso que toma en cuenta un modelo lineal asistido y la información auxiliar de la población proyectada por el DANE, con el censo de población y vivienda 2018. La metodología de modelos asistidos se puede revisar en Särndal, et al. (1992). El propósito de este documento es proporcionar la especificación técnica del modelo y su aplicación particular.

Para proceder a calcular los pesos calibrados, se define la matriz de diseño del modelo, donde cada columna identifica cada una de las viviendas, mientras que las filas representan los grupos de edad y sexo. De esta manera, cada celda representa el número de personas que conforman una vivienda en particular, según un grupo de edad y sexo específico. Toda vez que existen 4 grupos de edad y dos de sexo, entonces se tienen 8 filas. Adicionalmente, se crea otra fila con el número de hogares en la vivienda. En resumen, el número de columnas de la matriz es el número de viviendas encuestadas y el número de filas equivale a 9, que es el resultado de la combinación de las categorías de sexo y edad más el número de hogares. A continuación, se definen las categorías asociadas a cada variable de clasificación, para cabeceras y resto los grupos tienen una diferencia. A continuación, se presenta la tabla con la estructura.

Sexo:

- Hombres.
- Mujeres.

Grupos de edad:

En cabeceras:

- De 0 años a menores de 12 años.
- 12 años a menores de 25.
- 25 años a menores de 55.
- 55 años y más.

En resto:

- De 0 años a menores de 10 años.
- 10 años a menores de 25.
- 25 años a menores de 55.
- 55 años y más.

Como ejemplo se muestra la siguiente fila transpuesta de la matriz de diseño para una vivienda con 5 personas y 2 hogares.

Vivienda	Hombres < de 12 años	Mujeres < de 12 años	Hombres entre [12 -24) años	Mujeres entre [12 -24) años	Hombres entre [25 -54) años	Mujeres entre [25 -54) años	Hombres de 55 y + años	Mujeres de 55 y + años	Número de hogares
K	1	0	2	0	0	0	2	0	2

Fuente: Departamento Administrativo Nacional de Estadística

Entonces, siendo m el número de viviendas y 9 el número de combinaciones según grupos de sexo y edad más el número de hogares, cuya combinación se identifica con el sub-índice i , la matriz que contiene la información de individuos de la muestra tiene la siguiente forma:

$$X = \begin{bmatrix} x_{11} & \cdots & x_{1m} \\ \vdots & \ddots & \vdots \\ x_{91} & \cdots & x_{9m} \end{bmatrix}$$

De lo cual se tiene el vector de la vivienda k :

$$x'_k = [x_{1k}, x_{2k} \cdots x_{ik} \cdots x_{9k}]$$

Las proyecciones de la población son la información auxiliar que se usa para calibrar. Estas proyecciones entran como un vector que contiene, en sus celdas, las proyecciones de población de los diferentes grupos de edad y sexo predefinidos arriba y la proyección de hogares.

La información se pone en un vector de longitud 9:

$$t'_x = [t_1, t_2 \cdots t_i \cdots t_9]$$

De igual forma, se construye un vector con la estimación de la muestra de los 8 grupos de edad y sexo, más la estimación de hogares:

$$\hat{t}'_x = [\hat{t}_{x1}, \hat{t}_{x2} \cdots \hat{t}_{xi} \cdots \hat{t}_{x9}]$$

Donde $\hat{t}_{xi} = \sum_s d_k x_{ik}$ es una estimación del total de población de cada uno de los nueve grupos identificados con el subíndice i , s conjunto de viviendas de la muestra y d_k factor básico.

Siendo $\hat{T} = \sum_s d_k x_k x'_k$ una matriz 9×9 , se definen los pesos g_k como sigue:

$$g_k = 1 + (t_x - \hat{t}_x)' \hat{T}^{-1} x_k$$

Con todo el desarrollo anterior, se construye el factor de calibración:

$$w_k = d_k g_k$$

Nótese que los w_k son iguales para todos los individuos de la vivienda y diferentes entre viviendas, así mismo, se pueden construir los w_k con la metodología de estimadores de calibración (Särndal, et al., 1992), donde se demuestra que el estimador de calibración es asintóticamente equivalente al estimador de regresión bajo la métrica de la ji-cuadrado.

Para terminar la especificación técnica de los factores de expansión de la encuesta, se aclara que el modelo de calibración presentado no se aplica al conjunto total de la población en un solo modelo, sino que se desarrolla un modelo independiente por particiones geográficas. Las particiones donde se define cada modelo se conocen como bloques. No necesariamente los bloques deben coincidir con la estratificación.

Estimadores

Los estimadores son funciones con las cuales se estiman parámetros poblacionales. Para la encuesta se presentan los estimadores de totales y las razones con sus respectivos estimadores de varianza. Adicionalmente, con los anteriores se presentan los estimadores de los errores muestrales.

Estimador de Totales

El estimador del total para la variable y con la muestra s en el estrato h tiene la siguiente forma:

$$\hat{t}_{y_h} = \sum_{s_h} w_l y_l$$

Con l el sub-índice que representa al individuo.

El estimador del total poblacional de la variable y es entonces:

$$\hat{t}_y = \sum_{h=1}^H \hat{t}_{y_h}$$

En los casos en los cuales se requiera estimar grupos especiales de la población (dominios de estudio), la estrategia es crear otra variable que marque con 1 los individuos que pertenecen al grupo y cero los que no. Posteriormente, esta variable se multiplica a la variable y creando una nueva variable, llamémosla y^* . En este punto se procede de la misma manera para estimar.

$$\hat{t}_{y^*} = \sum_{h=1}^H \hat{t}_{y_h^*}$$

Cuando se quiere estimar unidades diferentes a personas, como es el caso de hogares o viviendas, la estrategia es similar a la de dominio de estudio y consiste en crear una variable con un uno para la

primera persona de cada una de estas unidades más grandes y cero para las demás personas dentro de la misma unidad.

Estimador de varianza del estimador del total

La encuesta no tiene un estimador de varianza exacto, teniendo en cuenta que es un diseño muestral complejo, en el cual en su última etapa se hace un sistemático de único arranque. Por lo tanto, el acercamiento a la varianza se plantea a través del método de conglomerado último (Hansen, et al. 1953).

$$\hat{v}(\hat{t}_y) = \sum_{h=1}^H \left(\left(\frac{n_h}{n_h - 1} \right) \sum_{i=1}^{n_h} \left(\hat{t}_{\hat{e}_{hi}} - \frac{\hat{t}_{\hat{e}_h}}{n_h} \right)^2 \right)$$

De donde i es el índice asociado a cada segmento de la muestra y toma valores desde 1 hasta n_h , el cuál es el número de segmentos dentro del estrato h . En algunos casos se adiciona a la formula el factor de finitud $(1 - N_h/n_h)$.

Si l es el índice que identifica cada individuo, k es el índice que identifica cada vivienda dentro del segmento i y l_i el número de individuos que pertenecen al segmento i , entonces:

$$\hat{t}_{\hat{e}_{hi}} = \sum_{l=1}^{l_{hi}} w_l \hat{e}_l$$

N_h el número de segmentos en el estrato h .

$$\hat{t}_{\hat{e}_h} = \sum_{i=1}^{n_h} \sum_{l=1}^{l_{hi}} w_l \hat{e}_l$$

$$\hat{e}_l = y_k - X'_k \hat{B}$$

$$\hat{B} = \hat{T}^{-1} \hat{t}$$

$$\hat{T} = \sum_{k=1}^K d_k x_k x'_k$$

$$\hat{t} = \sum_{k=1}^K d_k x_k y_k$$

Con y_k el agregado de individuos de la variable y en la vivienda k y K número de viviendas en el bloque donde se hace el modelo.

Estimador de Razones

El estimador de razón se construye haciendo el cociente de los estimadores de dos totales. Es así que si se quiere calcular la razón de los totales t_y y t_z se tiene la siguiente formula:

$$\hat{R} = \frac{\hat{t}_y}{\hat{t}_z}$$

Estimador de varianza del estimador de razones

El estimador de varianza de una razón es el mismo que el de estimador del total sólo que el total que se toma en cuenta es una variable producto de una transformación lineal. Más exactamente se usa el método de linealización de Taylor para poner a \hat{R} como combinación lineal de totales construyendo la variable artificial u mediante la expresión $u_l = \frac{1}{\hat{t}_z} (y_l - z_l \hat{R})$ sub-índice que identifica al individuo.

Entonces la varianza de la razón se puede calcular como:

$$\hat{v}(\hat{R}) = \hat{v}(\hat{t}_u)$$

Errores muestrales

Son diversas las formas en las cuales se puede medir la precisión de los indicadores. En la encuesta se presentan los coeficientes de variación estimados (errores relativos) y los errores marginales de cada indicador. Los estimadores de total y razón se asumen con una distribución normal y sus medias y varianzas son las presentadas en la anterior sección.

El coeficiente de variación se define como la relación porcentual del error estándar del estimador y el valor esperado del estimador. Se puede estimar de la siguiente forma:

$$CVe(\hat{t}) = \frac{\sqrt{\hat{v}(\hat{t})}}{\hat{t}} \quad \text{Coeficiente de Variación estimado para totales.}$$
$$CVe(\hat{R}) = \frac{\sqrt{\hat{v}(\hat{R})}}{\hat{R}} \quad \text{Coeficiente de Variación estimado para razones.}$$

El valor de este coeficiente, expresado en porcentaje, permite evaluar la calidad de un procedimiento de estimación. Se consideran como excelentes las mediciones menores de 5%, buenas las que están entre 5% y 10% y aceptables las que están entre 10% y 15%.

El error marginal mide la amplitud media del intervalo de confianza. Con este valor se puede definir el límite inferior y superior del intervalo. Para una confianza del 95% la fórmula de la estimación es:

$$EMe(\hat{t}) = 1.96^2 \sqrt{\hat{v}(\hat{t})} \text{ Error Marginal estimado para totales.}$$

$$EMe(\hat{R}) = 1.96^2 \sqrt{\hat{v}(\hat{R})} \text{ Error Marginal estimado para razones.}$$

2.2.13. Especificaciones de ponderadores

El software utilizado para el tratamiento de los ponderadores es un desarrollo propio del DANE en el programa SAS (Statistical Analysis System). Los resultados muestrales se desarrollan conforme a un ajuste por variable exógena, que es un estimador independiente de población, y permite mejorar las estimaciones referentes a las desagregaciones geográficas; se igualan los totales estimados con base en la muestra, con los totales del censo de población proyectados a la fecha de la encuesta, en este caso el CNPV de 2018. Lo anterior se apoya en la premisa de que la estructura por desagregaciones geográfica, proyectada a partir de un censo de población reciente, es más exacta que la estimada a partir de la muestra. A continuación, se muestra las etapas en las que se construyen los ponderadores:

- 1. Factores de corrección de cobertura:** Se calcula la cobertura de la muestra donde se identifican los segmentos y hogares perdidos en campo. Luego con la información de los conteos se construyen los factores de corrección en la base.
- 2. Factores básicos:** Se adiciona a la base los factores de expansión del diseño muestral. Se multiplican con los factores de corrección de la etapa anterior y como resultado se tiene un factor básico.
- 3. Factores techo:** Se toma en cuenta la población de cada estrato de muestreo para calibrar el factor anterior a las poblaciones censales del año 2018 de los estratos de muestreo preestablecidos en el diseño.
- 4. Factores de expansión calibrados:** Se usan las proyecciones poblacionales publicadas por el DANE, para ajustar según el modelo de calibración expuesto anteriormente el factor de techo.

El cálculo del ajuste por calibración se realiza utilizando la macro Clan 97 v3.1. (creada por Statistics Sweden), la cual corresponde a un conjunto de rutinas en el programa SAS para la estimación puntual y de errores en encuestas por muestreo.

5. Finalmente, a nivel del software STATA, se utilizan los objetos de base de datos, que se construyeron en el ítem anterior, y se realiza la programación para el cálculo de variables auxiliares (porcentajes, máximos y mínimos) y la generación del formato de salida. Una vez los cuadros se generan pasan al equipo de muestras para calcular los coeficientes de variación.

2.3. DISEÑO DE LA RECOLECCIÓN/ ACOPIO

Con la finalidad de garantizar un óptimo desarrollo del operativo de campo se desarrolla el diseño de la recolección en ocho partes: 1) Método y estrategias de recolección de datos; 2) Estructura organizacional del operativo; 3) Esquema de entrenamiento; 4) Conformación del equipo; 5) Proceso de sensibilización y acuerdos de intercambio; 6) Elaboración de manuales; 7) Diseño de las estrategias de comunicación y 8) Diseño de estrategias de seguimiento y control.

2.3.1. Métodos y estrategias de recolección de datos

El método de recolección de la información de la encuesta es la entrevista directa, la misma se desarrolla con cada una de las personas que conforman los hogares que sean informantes idóneos⁶ de las viviendas seleccionadas.

Por su parte, el sistema de recolección es el barrido, el cual consiste en que cada equipo de trabajo recorre simultáneamente el área asignada, hasta investigar todas las unidades de vivienda que la conforman.

El medio de recolección empleado es el Dispositivo Móvil de Captura (DMC). Cada encuestador dispondrá de un DMC, tarjeta de memoria extraíble (SD), cargador portátil, tarjeta SIM, software de captura correspondiente a la encuesta y formularios en papel para aquellos casos en los que no sea posible hacer uso del DMC.

Transmisión de datos

El proceso es realizado por el Gestor informático de la sede, quien es la persona encargada de realizar el descargue de la información almacenada en los Dispositivo Móviles de Captura (DMC) y la transmisión de la información que esté completa y haya sido capturada y depurada durante el operativo de campo hacia la Oficina de Sistemas de DANE Central. Esta transmisión se efectúa a través del protocolo de transferencia de archivos FTP, previamente establecido por la oficina de Sistemas de DANE Central.

Previo a la transmisión de los puntos de las encuestas al DANE Central, el Gestor informático comprueba que la información que se cargue en el FTP sea consistente con la que posee el Coordinador de campo en el resumen de cobertura.

2.3.2. Estructura organizacional del operativo y conformación del equipo

Esquema operativo

Los equipos de trabajo en cada sede se conforman según el tamaño de la muestra y las características principales de cada departamento. Cada equipo operativo en campo se conforma por un supervisor y puede tener entre dos y tres encuestadores. Esta organización varía de acuerdo con la distribución de las cargas de trabajo.

Organigrama

El desarrollo del trabajo de campo es responsabilidad de las direcciones territoriales y de la coordinación logística del DANE Central. A continuación, se presentan de manera general las actividades para cada uno de los roles que participan en el operativo y el organigrama del equipo de trabajo, en cada una de las direcciones territoriales:

Asistente de encuesta: es la persona responsable de las actividades operativas preliminares de la encuesta, así como de la supervisión en terreno a los equipos de trabajo. También tiene a su cargo el control de calidad de los datos e información recolectada y el manejo operativo. Generalmente, en las diferentes sedes del DANE esta responsabilidad la asume una persona de planta.

Coordinador de campo: en coordinación con la asistencia de encuesta, son los responsables de las actividades operativas y la supervisión de los equipos de trabajo en campo, incluyendo la crítica y digitación de los datos recolectados. De igual forma, tiene a su cargo el control de la calidad de la información y el manejo de la encuesta en terreno y debe asegurar que toda la información operativa requerida llegue oportunamente.

Supervisor de campo: tiene por objeto asegurar la calidad de la información recolectada y la cobertura en los hogares y personas seleccionadas en la muestra. Así mismo, coordina, dirige y controla el operativo en los segmentos que se le asignan, reportando su trabajo a la Coordinación de la encuesta, también verifica el cumplimiento de las instrucciones metodológicas y operativas del trabajo en campo de las personas que desarrollan las actividades de recolección. Tiene a su cargo mínimo 2 y máximo 3 recolectores.

Sensibilizador: se encarga de dar a conocer con anticipación a los hogares el contenido y propósito de la encuesta, mediante el barrido de cada segmento de la muestra, a través de piezas comunicativas dirigidas a las personas y hogares seleccionados y objeto de observación.

Recolector: se encarga de obtener la información requerida por la encuesta de forma presencial, mediante entrevista directa en cada uno de los hogares y segmentos asignados por la Supervisión, en cumplimiento de las normas y conceptos establecidos en el manual de recolección y conceptos básicos. Reporta su trabajo directamente a la Supervisión de Campo.

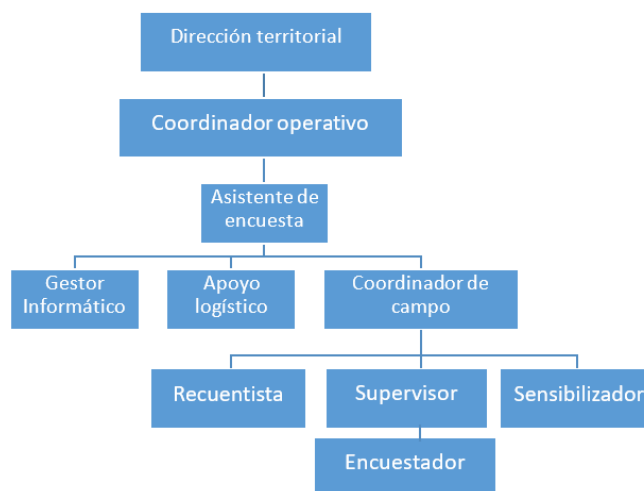
Esta tarea es fundamental, toda vez que debe asegurar la calidad y el correcto diligenciamiento de la encuesta a todos los hogares definidos en la muestra y que le hayan sido asignados por quien ejerce la supervisión, su labor consiste en realizar las visitas a los hogares cuantas veces sea necesario para cumplir a cabalidad con la captura de la información requerida.

Recuentista: realiza el recorrido de un área previamente asignada y efectúa el conteo de las unidades y viviendas que la conforman, con el fin de generar la correcta asignación de medidas de tamaño (MT) y selección de los segmentos a encuestar.

Gestor informático: es responsable de la actualización de los aplicativos de captura en su sede y de la transmisión periódica de la información recolectada, así como la de realizar el cruce de segmentos faltantes con lo transmitido. Junto con la Coordinación de campo verifica que la información que se cargue en el FTP sea consistente con el resumen de cobertura.

Apoyo logístico: es el encargado de apoyar la elaboración de informes, formatos y demás documentación que se requieran para cumplir con las labores y tareas previstas antes, durante y con posterioridad al operativo de campo, incluyendo las actividades de cierre de este.

Figura 3. Organigrama del equipo de trabajo



Fuente: Departamento Administrativo Nacional de Estadística

En el DANE Central los aspectos metodológicos y muestrales están a cargo de los equipos de Temática y Diseños Muestrales respectivamente. Igualmente, los aspectos logísticos y operativos le corresponden al grupo de la Dirección de Recolección y Acopio. El desarrollo del operativo será responsabilidad de cada Dirección territorial y las labores de procesamiento de datos las realizará la Oficina de Sistemas del DANE Central.

2.3.3. Esquema de entrenamiento de personal

Se realizará el entrenamiento al personal que cumpla con los requisitos mínimos para la conformación del Banco de Prestadores de Servicios Operativos (BPSO), ya sea de forma virtual o presencial. De igual forma, se contará con los formatos de las evaluaciones a aplicar, las listas clasificadas por rol, al igual que los listados con los resultados de las evaluaciones aplicadas posteriores al aprendizaje.

Con el fin de garantizar el adecuado entrenamiento del personal y obtener la mejor cobertura y calidad de la información en la ejecución del operativo de campo, se realizarán sesiones de aprendizaje de la siguiente manera:

Entrenamiento primer nivel:

Se realizará el entrenamiento de aprendizaje virtual a través de la plataforma AprendeNET conforme a los lineamientos establecidos en el modelo B-Learning. Los aspirantes tendrán acceso a los módulos diseñados por cada una de las áreas del DANE central (Dirección de Recolección y Acopio, Oficina de Sistemas, Dirección de Geoestadística, Dirección de Metodología y Producción Estadística, Dirección de Difusión y Cultura Estadística) en donde se darán a conocer los temas a tratar durante el desarrollo del operativo. Mediante la metodología usada se logra interactuar de manera dinámica con el material de estudio, facilitando y motivando el proceso de aprendizaje en la etapa virtual.

Entrenamiento segundo nivel:

Desde el GIT de Encuesta a personas e instituciones de la Dirección de Recolección y Acopio del DANE central se organizará el reentrenamiento operativo de forma virtual mediante la plataforma Microsoft Teams, para las sedes participantes en la ENTIC hogares. En este entrenamiento participarán las áreas de Dirección de Recolección y acopio, Sistemas, Dirección de Geoestadística, Dirección de Metodología y Producción Estadística y Dirección de Difusión y Cultura Estadística.

Lo ideal es contar con la presencia de todas las personas contratadas para los roles de Asistente de encuesta, Coordinador de campo, Supervisor, Recolector, Sensibilizador, Recuentista y Gestor informático.

2.3.4. Conformación del equipo

La convocatoria y la selección del personal, se realiza siguiendo los lineamientos establecidos por la Secretaría General del DANE Central.

En conjunto con las Direcciones territoriales, la Dirección de Recolección y Acopio de DANE Central define los perfiles, cantidad de vacantes, tiempos y el presupuesto asignado para el personal a contratar. Teniendo en cuenta dichos perfiles, se procede a abrir la invitación pública en la página web del DANE.

El proceso de selección inicia con la inscripción en el banco de hojas de vida de la entidad, donde los aspirantes podrán registrar su formación académica y experiencia laboral, para poder formar parte del proceso de selección del personal operativo requerido para los diferentes proyectos que adelanta la entidad a nivel nacional.

2.3.5. Proceso de sensibilización y acuerdos de intercambio

La sensibilización tiene por objeto lograr una respuesta oportuna y de calidad por parte de las fuentes de información. Se busca propiciar un ambiente de colaboración, acercamiento y conocimiento de la población objetivo mediante la visita a cada uno de los hogares a los que les será aplicada la encuesta.

En esta primera visita, se informan las fechas previstas en las cuales el grupo de recolección del DANE puede aplicar la encuesta. En caso de que la sensibilización no encuentre a los miembros de los hogares en la primera visita, se realizan revisitas con el fin de sensibilizar a las personas que se encuentren en las viviendas y hogares asignados en la muestra.

La sensibilización, en las ciudades capitales y áreas metropolitanas es llevada a cabo por una persona contratada para tal fin. En el caso de la recolección rural que se realiza en centros poblados y rural disperso, la sensibilización es llevada a cabo en paralelo con el proceso de recolección y es asumida por el supervisor de campo y su equipo operativo.

La investigación cuenta con herramientas para realizar el proceso de sensibilización como lo son los folletos informativos y las notificaciones de la encuesta que se entregan en cada uno de los hogares seleccionados en la muestra.

2.3.6. Elaboración de manuales

Para el control de los procesos operativos se cuenta con los siguientes manuales y formatos:

Plan de recolección: usado como lineamiento para los diferentes aspectos que se deben tener en cuenta para el desarrollo del operativo de campo en materia de procesos pre-operativos y operativos, recursos, métodos, sistemas e instrumentos de recolección, entre otros.

Manual del coordinador de campo: describe las actividades y consideraciones a tener en cuenta en la coordinación en campo, ofreciendo una guía de las acciones a realizar en este aspecto.

Manual del supervisor: utilizado para suministrar la información que el supervisor de campo requiere junto con las normas establecidas e instrucciones necesarias que debe seguir y aplicar en el cumplimiento de los objetivos propuestos para el operativo.

Manual de recuento: se tiene en cuenta con el fin de otorgar al recuentista o supervisor rural las pautas para el desarrollo de las actividades específicas del proceso de recuentos de las unidades.

Manual de diligenciamiento de formatos: hace uso de este el personal en todos los roles, pues en él se presenta de manera clara y precisa el diligenciamiento de los formatos tipo registro que deben ser utilizados por el personal de campo.

Componente temático

Manual de recolección: usado como guía durante el operativo para resolver dudas e inquietudes a lo largo de la encuesta. Cuenta con una explicación detallada y la forma de resolver cada una de las preguntas de la encuesta. Contiene la definición de los conceptos básicos e incluye ejemplos que ayudan a entender mejor los temas de la encuesta.

Componente sistemas

Manual del usuario: este manual contiene la descripción de los procedimientos utilizados para el diligenciamiento de información en el aplicativo de captura, la realización de copias de seguridad y el cierre de encuestas, facilitando al usuario la recolección de la información y el manejo de cada una de las preguntas de la encuesta.

2.3.7. Diseño de las estrategias de comunicación y plan de contingencias

Como estrategia de comunicación, se desarrolla un proceso de sensibilización con enfoque de relacionamiento en las tres fases del proceso, las cuales son el preoperativo, el operativo y el posoperativo, dentro del cual se busca:

- Dar a conocer la importancia del DANE y de la encuesta.
- Identificar las entidades gubernamentales, autoridades locales, agentes externos y líderes comunitarios, logrando establecer comunicación y apoyo en el relacionamiento con la comunidad donde se aplica la recolección.
- Sensibilizar a cada integrante del hogar, estableciendo comunicación directa con todos los miembros e iniciando con el jefe de hogar.
- Reportar las novedades presentadas que dificulten el acceso a las fuentes, como por ejemplo rechazos, restricción de acceso a conjuntos residenciales, territorios étnicos o de otros grupos de la comunidad.
- Mantener un relacionamiento adecuado con las fuentes durante la recolección de información dejando “las puertas abiertas” al DANE para próximas encuestas.

La recepción a inquietudes, novedades u otras eventualidades presentadas en el operativo ENTIC Hogares se realizará por medio de correo electrónico, a éstos se le hará el seguimiento correspondiente para brindar respuesta oportuna.

2.3.8. Diseño de la estrategia de seguimiento y control

Para el control de calidad de la información se utilizan normas de validación y consistencia y controles operativos de calidad tales como el resumen de cobertura, las alarmas e inconsistencias, el monitoreo, los indicadores de cobertura de vivienda, hogares y personas y el Indicador de Calidad de la Recolección (ICR). Estos serán explicados a continuación:

- **Especificaciones de validación y consistencia**

En primera instancia, el grupo temático define, con base en las características de la investigación, una serie de validaciones y consistencias entre preguntas, filtros y flujos del formulario. Éstas son implementadas en el aplicativo dispuesto en el DMC por la Oficina de Sistemas, junto con validaciones adicionales definidas por el grupo de la DRA para el seguimiento y control de los resultados operativos.

Este proceso se efectúa mediante pruebas directas en las que participan los equipos que hacen parte de la encuesta (DRA, Sistemas, Diseños muestrales y Temática). Se destaca que esta etapa se realiza antes de iniciar el operativo en campo, dado que el propósito es hallar y corregir posibles errores o dificultades que contenga el formulario de la encuesta.

- **Instrumentos de control para supervisión**

Para el proceso de control, el grupo de la Dirección de Recolección y Acopio de la encuesta diseñó un conjunto de formatos que facilitan el monitoreo y supervisión tanto del proceso de recuento, como de sensibilización y recolección de la información. Esto permite comparar los datos obtenidos en campo con el reporte generado en la base de datos de encuestas transmitidas. Entre ellos se enmarcan:

- Resumen acumulado de viviendas, hogares y personas por segmento.
- Resumen de cobertura.
- Formato diario de supervisión de campo.
- Informe técnico de coordinación de campo.

Las sedes son responsables de verificar y asegurar la calidad de la información recolectada y transmitida al DANE Central. El seguimiento a la calidad inicia con el supervisor de campo, quien hace la respectiva verificación confirmando que todas las encuestas estén totalmente diligenciadas y con la calidad establecida, de modo que esta es una actividad que hace al momento de cerrar la encuesta.

Ahora bien, desde DANE Central se hará verificación de la información durante el operativo para que las sedes tomen correctivos oportunamente a cada una de las inconsistencias. Semanalmente se generan reportes de cobertura, los cuales son enviados a los asistentes de encuesta para que verifiquen el estado de cobertura del operativo, además de observar los avances en la cobertura estimada.

Resumen de cobertura

El objetivo de este formato es llevar un control de manera resumida del número de viviendas, hogares y personas encontradas en los segmentos asignados y trabajados durante la recolección para detectar posibles inconsistencias y alarmas en la información recolectada. Este formato es utilizado en el operativo urbano, en las cabeceras municipales y en el área rural disperso.

El principal usuario del formato es la persona encargada de la coordinación de campo, quien consolida la información de la supervisión a cargo para remitir semanalmente estos resúmenes a la Dirección de Recolección y Acopio en el DANE central vía FTP.

De este reporte se obtienen los indicadores de cobertura operativa y resultados que permiten medir el proceso de seguimiento.

Detección de alarmas y corrección de inconsistencias

Del resumen operativo de cobertura se obtienen las inconsistencias detectadas en la revisión semanal. Al realizar la revisión, se reportan las cifras que son motivo de justificación por parte de las sedes tales como la baja cobertura, segmentos con menos de 10 viviendas encontradas, alto número de vacantes, entre otros. Las inconsistencias deben ser corregidas por la ciudad correspondiente, así mismo las alarmas deben ser justificadas en el menor tiempo posible.

Identificación de errores temáticos y solicitud de revisión

Además de las inconsistencias revisadas durante el operativo de campo por el equipo de la DRA, en el proceso de carga de las bases de datos el equipo temático identifica posibles errores de diligenciamiento, enviando reportes a las territoriales para verificar si son errores efectivos o si existe una justificación para las alertas generadas.

Monitoreo

Periódicamente se realizan monitoreos a los grupos de campo para observar el desempeño de éstos y detectar falencias que pueden afectar la calidad de la información recolectada. En este orden, se observan aspectos como la presentación personal; la lectura adecuada de las preguntas; la no-inducción de respuestas; los cambios de capítulo; la realización de buenos sondeos; el registro adecuado de los miembros del hogar y el manejo de conceptos. De igual forma, se verifica la correcta distribución de materiales de trabajo, el diligenciamiento de formatos, el manejo del transporte, la revisión de las encuestas por parte del personal de supervisión, la ubicación en el segmento y la capacidad para crear un ambiente propicio con la fuente que permita desarrollarla.

Todos los aspectos para mejorar son socializados con los responsables de cada proceso y se hace seguimiento a la puesta en marcha de las soluciones y correctivos que sean necesarios.

Informe de cobertura campo vs sistemas

Semanalmente, luego de consolidar la totalidad de la información de cobertura de campo de cada una de las sedes, se efectúa un cruce de variables entre la base de sistemas y el resumen operativo de cobertura, comparando la cantidad de viviendas, hogares y personas a nivel de segmento.

La información consolidada y comparada se envía a las sedes para su respectiva revisión y corrección. De esta forma, las sedes pueden detectar información con errores de digitación en los resúmenes de cobertura y la posible falta de información o duplicidad de viviendas en la base de sistemas.

A partir de esto, las ciudades deben hacer los ajustes respectivos al resumen de cobertura y transmitirlo nuevamente vía FTP. Si la información de campo es correcta, deben verificar las inconsistencias en la

base de sistemas y enviar nuevamente los archivos respectivos comunicando la novedad al ingeniero encargado en el DANE Central.

Indicadores de cobertura en viviendas, hogares y personas

Semanalmente se realiza seguimiento a las tasas de cobertura operativa, con el fin de detectar oportunamente comportamientos atípicos que permitan presumir debilidades en el desempeño de los grupos de trabajo en campo, y así efectuar indagaciones más profundas para tomar los correctivos respectivos.

- **Cobertura de viviendas:** se obtiene de cruzar la información de las viviendas finales encontradas en el operativo de campo contra el total de viviendas encontradas en sensibilización * 100. Los datos son obtenidos de lo reportado en el resumen de cobertura y de lo encontrado en sensibilización.
- **Cobertura de hogares:** se obtiene de cruzar la información de hogares con encuesta completa sobre cantidad de hogares obtenidos en sensibilización * 100. Los datos son obtenidos de lo reportado en el resumen de cobertura y de lo encontrado en sensibilización.
- **Cobertura de personas:** se obtiene de cruzar la información del total de personas con encuesta completa contra la cantidad de personas obtenidas en sensibilización * 100. Los datos son obtenidos de lo reportado en el resumen de cobertura y de lo encontrado en sensibilización.

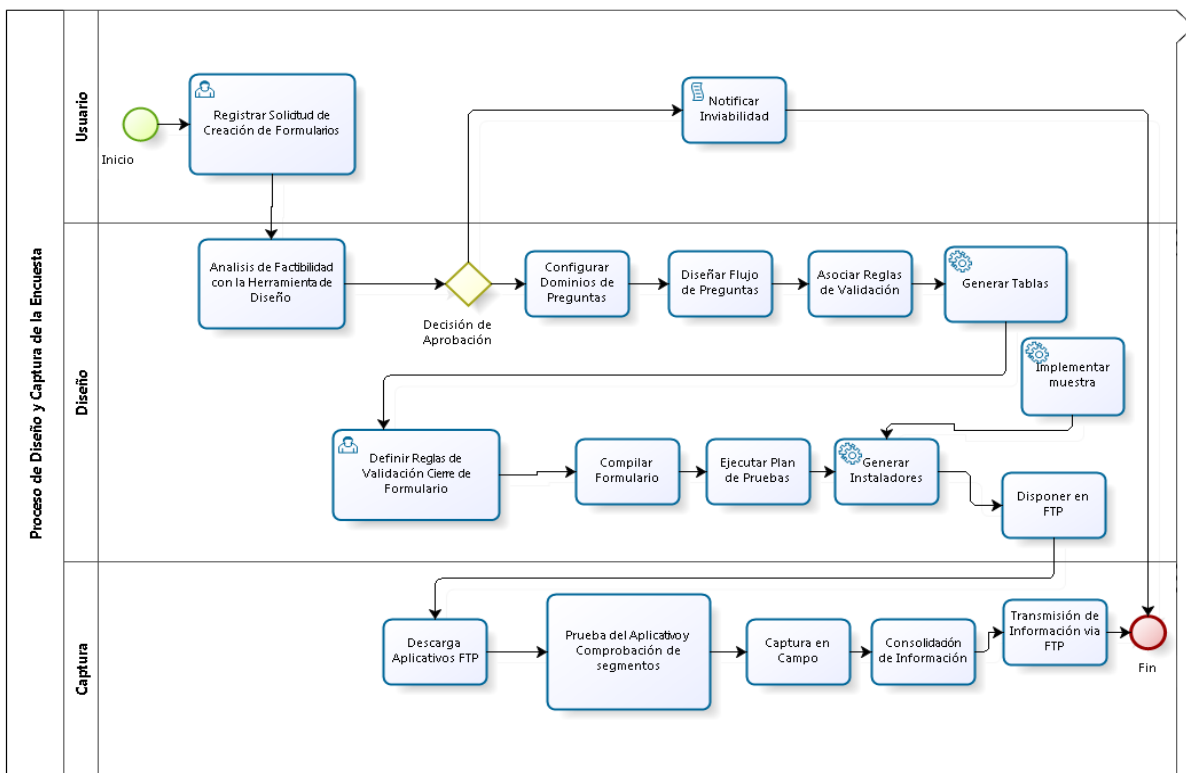
Indicador de Calidad de la Recolección (ICR)

El objetivo de este indicador es determinar el grado de calidad y confiabilidad de la información enviada por la persona encargada de la coordinación de campo de cada sede. Este indicador se mide a través del número de inconsistencias detectadas en el acompañamiento al personal encuestador en campo. Esto permite hacer seguimiento continuo a las diferentes sedes que presentan bajos indicadores de calidad de la información, con el fin de determinar las posibles fallas y sugerir recomendaciones o medidas de mejoramiento.

Se calcula un indicador de calidad para cada persona del grupo de recolección, el cual es obtenido con el promedio aritmético de los puntajes alcanzados para los formularios que fueron sometidos al proceso de supervisión y control de calidad, mediante un formato dispuesto especialmente para este fin. Dentro del cálculo del tipo de error cometido se crea una variable de penalización sobre el resultado del indicador para hacer distinción de dichas faltas, las cuales se penalizan de acuerdo con el tipo de inconsistencia.

2.3.9. Diseño de sistemas para la obtención de datos

Figura 7. Proceso de Diseño de sistemas para la obtención de datos



Fuente: Departamento Administrativo Nacional de Estadística

Elementos de Proceso



Inicio

Inicio del proceso de solicitud de generación de formulario, encuesta o investigación.



Registrar Solicitud de Creación de Formularios

Entrega de los documentos con los insumos para la encuesta por parte del componente temático de la investigación.



Análisis de Factibilidad con la Herramienta de Diseño

Se evalúa la viabilidad de implementación del formulario en el diseñador de encuestas, según la documentación entregada (Formulario y normas de validación y consistencia).



Decisión de Aprobación

Aprueba o Rechaza la solicitud



Configurar Dominios de Preguntas

En el diseñador de encuestas se implementan los dominios de preguntas de tipo numérico, texto, fecha, lista de valores (única, múltiple) o tablas prediseñadas en formato CSV.



Diseñar Flujo de Preguntas

El diseño del flujo de preguntas se organiza a través de la hoja de encuesta establecida para ello en el diseñador de encuestas.



Asociar Reglas de Validación

Se implementan las reglas de validación de acuerdo con el documento de especificaciones y consistencia; se validan entre otras, rangos, tipos de datos, flujo de datos, operaciones aritméticas.



Generar Tablas

La generación de tablas en formato CSV se realiza para aquellas listas que tienen gran cantidad de dominios, con el fin de hacer más ágil el llamado de estas dentro del formulario.



Definir Reglas de Validación Cierre de Formulario

Las reglas de validación de cierre de encuestas involucran el ingreso de preguntas de control de flujo de la encuesta, a nivel de vivienda, hogar, personas y capítulos específicos, en donde a través de especificaciones se controla el acceso a determinados capítulos del formulario.



Compilar Formulario

Tras elaborar el formulario se puede compilar este para validarlo y verificar que cumpla con las funcionalidades definidas en las especificaciones.



Ejecutar Plan de Pruebas

Una vez diseñados y modelados los procesos para la captura de información, se procede a realizar un plan de pruebas que permita evaluar el formulario y determinar su salida a campo, mediante pruebas de escritorio y consistencia de datos.



Implementar muestra

Permite cargar en el formulario la información de los segmentos y manzanas a recolectar definidos en la muestra.



Generar Instaladores

Los instaladores se generan por ciudad o municipio, según las especificaciones diseñadas por el componente logístico de la investigación.



Disposición y Descarga de Aplicativos en FTP

Los instaladores se disponen en el buzón FTP en formato comprimido y posteriormente descargado por el apoyo de sistemas responsable en cada ciudad o municipio, que tiene nombre de usuario y clave de autenticación para acceder al buzón.




Prueba del aplicativo y Comprobación de Segmentos

El apoyo informático de cada una de las ciudades realiza pruebas de escritorio, simulación en campo y se encarga de validar el formulario y la muestra allí implementada para la encuesta.




Captura en Campo

Una vez digitados los datos en los DMC (Dispositivos Móviles de Captura), se debe generar un backup de la información recolectada diariamente por cada persona encuestadora. El personal de supervisión debe validar y consistencia de la información para el cierre de cada encuesta; la información debe entregarse al apoyo informático.

 Consolidación de Información

Consiste en ir almacenando en el computador personal del apoyo de sistemas, en una estructura de archivos definida, todas las encuestas realizadas, para su posterior envío.

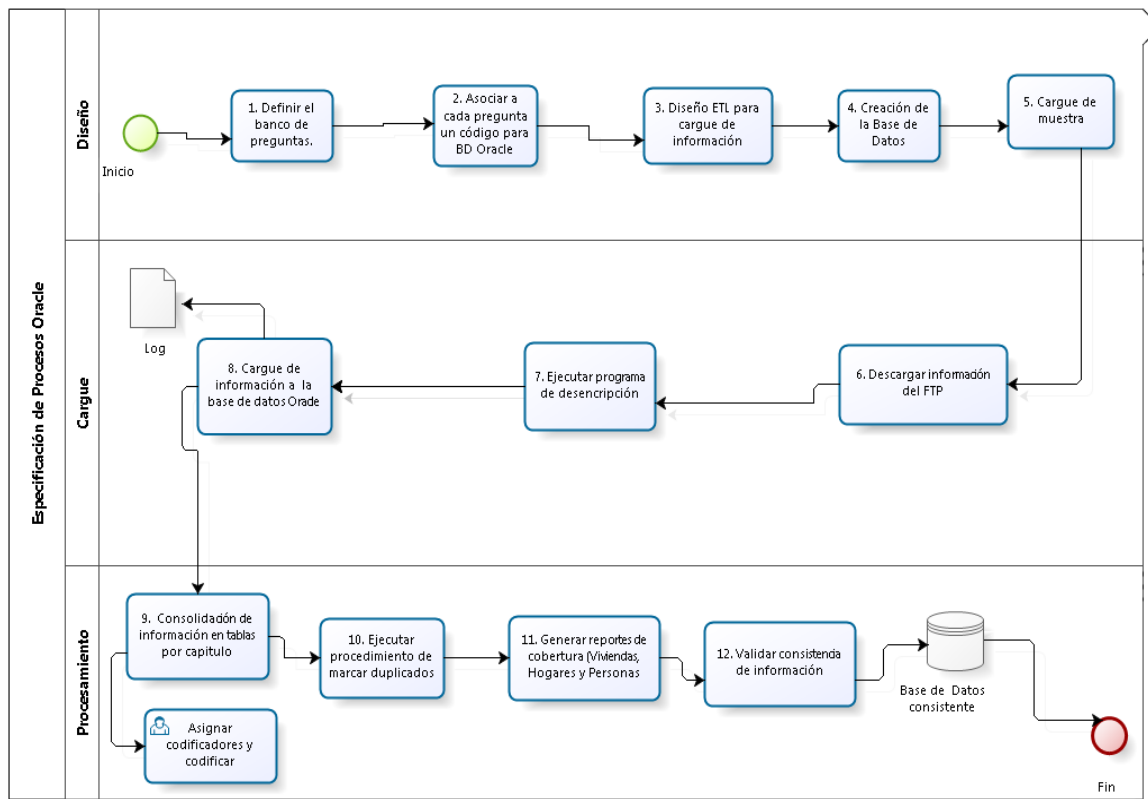
 Transmisión de Información vía FTP

Una vez capturada la información diariamente se comprime y se dispone en el buzón FTP creado, para su posterior análisis y procesamiento El coeficiente de variación.

 Fin

Finalización del proceso.

2.3.10. Transmisión de datos

Figura 8. Especificación de procesos


Fuente: Departamento Administrativo Nacional de Estadística.

- **Creación de la Base de Datos.**

En el motor de base de datos se crea un esquema para la investigación que permite el cargue y almacenamiento general de los datos. La base de datos para la investigación asignada es relacional, en la cual cada tabla corresponde a un capítulo de la encuesta y existe un único registro por vivienda.

En este sentido, se debe asignar el número de la Encuesta asociado a la investigación y definir los tipos de datos, texto de la pregunta, listas de valores asociadas, subpreguntas, entre otros. De igual manera, es fundamental definir los formularios y subformularios para las encuestas a hogares, en los cuales generalmente se incluyen los tres niveles básicos para la toma de información, los cuales son vivienda, hogar y personas. Cada nivel relacionado consta de uno o varios capítulos según la temática y opcionalmente se establecen los flujos y validaciones de la totalidad de las preguntas.

- **Envío de información**

Diariamente se envía información de encuestas completas al DANE Central. Para este proceso se consolida la información y se realiza la transmisión de datos de las encuestas a través del FTP (File Transfer Protocol), el cual es un protocolo de red para la transferencia de archivos entre sistemas conectados a una red, basado en la arquitectura cliente-servidor. Desde el equipo cliente las ciudades o municipios se conectan a un servidor para descargar archivos desde él o para enviarlos.

A cada ciudad o municipio se le crea un buzón en el servidor FTP del DANE Central, donde se informa de las encuestas completas recolectadas diariamente. Al final del operativo, se incluye toda la información faltante de los segmentos asignados para recolección (Encuestas completas e incompletas).

2.4 DISEÑO DE PROCESAMIENTO

A continuación, se presenta el diseño de las herramientas tecnológicas de software y hardware, que se implementan para el procesamiento de los datos (software y hardware), los programas requeridos para la grabación, la consolidación y el almacenamiento.

2.4.1. Consolidación de archivos de datos

Recibidos los archivos, la información es descargada del buzón FTP y almacenada en una estructura jerárquica que permite realizar la descrición de los datos. Una vez se seleccionan los archivos que se deben cargar en la base de datos, se realiza el proceso de carga utilizando herramientas de ETL (extracción, transformación y carga).

- **Cargue de información**

Al iniciar el proceso de carga se verifica la identificación única para cada encuesta, garantizando que no exista duplicidad en la información y se asigne un directorio único a cada encuesta.

Una vez cargados los datos, se genera un log que indica el estado o posibles fallas presentadas durante el proceso y se procede a trasponer toda la información. Este proceso permite poblar las tablas creadas en la base de datos para cada capítulo de la investigación. Así, cada tabla se relaciona con cada nivel manejado en la encuesta (vivienda - hogar - persona).

- **Consistencia de información**

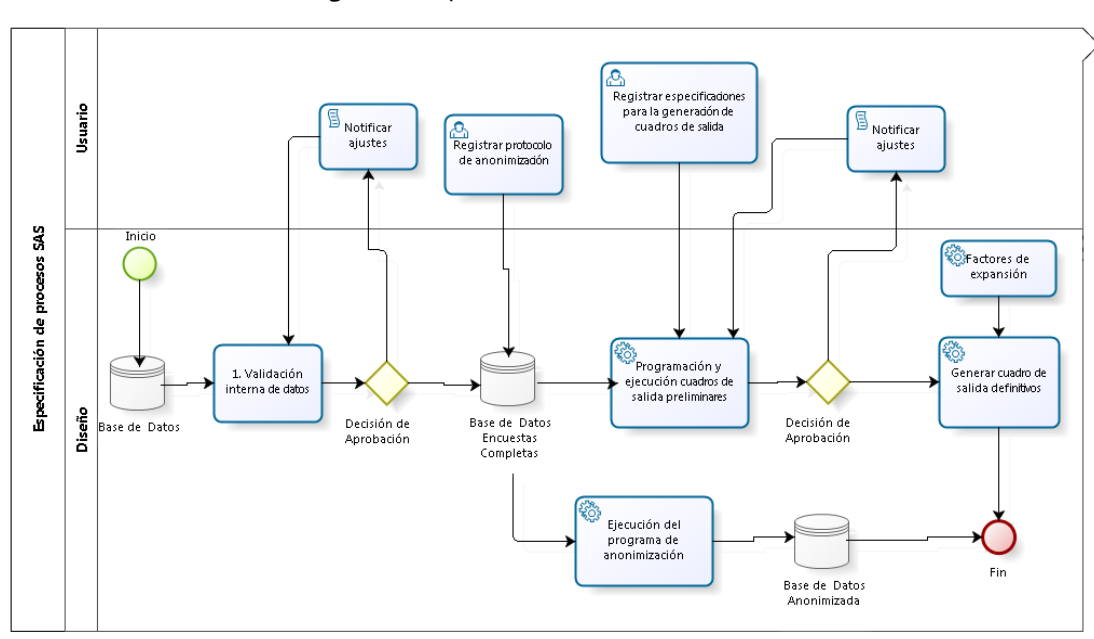
A través del desarrollo de sentencias DML (Data Manipulation Language, por sus siglas en inglés), se identifican posibles inconsistencias por duplicidad a nivel de personas y problemas de selección de segmentos a trabajar. Así mismo, se generan reportes de cobertura y frecuencias que permiten realizar constantemente el monitoreo y control a la información que se ha recolectado y cargado en las bases de datos.

Después, en el esquema, definido en el motor de base de datos, se crean procedimientos para validar la información de la base de datos, flujos, variables sin información, rangos y estructura, para obtener una base de datos consistente.

Así, cada investigación se almacena en su esquema de base de datos. Esto permite que los datos se mantengan aislados de la información de otras investigaciones y los permisos de acceso a los usuarios se realicen mediante los mecanismos de roles y privilegios propios del sistema manejador de base de datos.

Producción de información

Figura 9. Especificación de Procesos SAS



Fuente: Departamento Administrativo Nacional de Estadística.

2.4.2. Codificación

En la estructura del formulario se debe asociar a cada pregunta el código equivalente con el de pregunta en las tablas asignadas para cada capítulo en el esquema de datos, incluyendo el número de formulario. Adicionalmente, como parte del diseño se crea un ETL (Extracción Transformación y Carga, por sus siglas en inglés) y se carga la muestra a una tabla de la base de datos, para poder realizar el cargue de la información correspondiente.

2.4.3. Diccionario de datos

En el diccionario de datos se describen por temáticas las características de todas las variables que son (o serán) utilizadas en la operación estadística en concordancia con las tablas de datos generadas. Así mismo, este apartado ayuda a definir las especificaciones de estimación (para operaciones estadísticas por muestreo). Todo esto, se realiza con base el formulario de la encuesta realizada, que será actualizado cada año.

El diccionario de datos de la ENTIC hogares, facilita el cumplimiento de las siguientes tareas:

- **Configuración de la base de datos:** permite visualizar la estructura de la base de datos por capítulos, variables temáticas y/o por unidades de observación para todas las preguntas aplicadas con el instrumento de recolección en el operativo de campo. También, permite realizar consultas rápidas sobre la caracterización de variables, presentando la codificación de las variables, tipo, extensión y posibles respuestas.
- **Reglas de edición (validación y consistencia) e imputación:** permite realizar una revisión de las variables para su dominio geográfico en términos de viviendas, hogares y personas. Así mismo, es utilizado por el equipo de temática para realizar las respectivas pruebas de inconsistencias y ser informadas oportunamente para resolverse con el equipo de campo.

2.4.4. Revisión y validación

Las sedes realizan los procesos de revisión y verificación y luego se transmite la base de datos al DANE Central, en esta etapa, los analistas de la operación estadística realizan la revisión de la consistencia de la información. En el evento en que se detecten inconsistencias, las mismas se resuelven en la fase de devolución, en la cual, la sede debe realizar las respectivas observaciones. Estos procesos de aclaración

o corrección de la información se realizan por teléfono, correo electrónico o visitas a la fuente cuando la situación lo amerite.

Una vez la base de datos es aprobada por el equipo de la Dirección de Recolección y Acopio (DRA), la misma es remitida al equipo temático de la operación estadística, el cual realiza una revisión de calidad a través de un análisis de la información a nivel de agregados totales por departamento, cabecera municipal y centros poblados y rural disperso y agregados por sexo. Para ello, se recurre a calcular variaciones, contribuciones y participaciones; sin embargo, si tras el análisis se encuentran inconsistencias o aclaraciones sobre la información, se retorna al grupo de analistas del DANE Central, quienes deben aclarar las respectivas correcciones de la información, consultando directamente con las fuentes de información. Posterior al desarrollo de dicho proceso, la base de datos se entrega depurada al equipo de temática que luego del análisis y aprobación de las aclaraciones o ajustes realizados a la base, procede a la elaboración de los productos para publicación.

Posteriormente, se realiza un comité interno en el cual un grupo de funcionarios de las diferentes áreas internas del DANE (Dirección de Recolección y Acopio, la Dirección de Regulación, Planeación, Estandarización y Normalización, la Dirección de Metodología y Producción Estadística y la Dirección de Síntesis y Cuentas Nacionales) realizan el análisis de la información tanto a nivel de la información agregada como a nivel del contexto durante el período de referencia. De igual forma, el equipo de DIRPEN realiza una validación de las métricas de calidad para la iteración de la ENTIC Hogares que se piensa publicar.

2.4.5. Diseño de instrumentos de edición (validación y consistencia) e imputación de datos

En el procesamiento de datos se utiliza el programa informático SAS (Statistical Analysis System) diseñado para el análisis de datos y soluciones estadísticas. Para esto se lleva a cabo la verificación interna de los datos y ajustes; ajustes de cobertura y generación de cuadros de salida y el proceso de anonimización.

- **Verificación de la consistencia interna de los datos y ajustes**

Se realiza la revisión de totales y subtotales para cada variable dependiendo del dominio geográfico en términos de viviendas, hogares y personas. Así mismo, se generan las frecuencias de las principales variables y se realizan análisis de las respectivas distribuciones de los valores extremos y de los totales entre otros. Finalmente, como resultado de este proceso se realiza un reporte de inconsistencias que es enviado al equipo de temática, el cual solicita los respectivos ajustes y corrección de éstas cuando son necesarias.

- **Imputación y ajustes de cobertura**

Una vez cumplidas las etapas de consistencia, depuración, validación de la información y revisión de frecuencias y cumplida la totalidad de los procesos estadísticos que garantizan la calidad y cobertura de la investigación, se realiza la conformación de la base solamente con las encuestas completas. El componente temático de la encuesta hace entrega de los documentos con las especificaciones para la generación de cuadros de salida. Éstos son desarrollados y posteriormente se ejecutan los programas que permiten generarlos sobre la base de datos de encuestas completas y se efectúan los análisis de resultados, los cuales son generados en HTML o XLS.

Revisados los cuadros generados por parte del equipo temático, se realizan ajustes a los programas en caso de ser necesario, posteriormente se reciben los factores de expansión por parte del equipo de muestras y se generan los cuadros de salida definitivos, que permiten la visualización de los resultados de la encuesta.

2.4.6. Diseño para la generación de cuadros de resultados

El diseño de los cuadros de resultados describe la totalidad de los procesos desarrollados para el llenado de cuadro de salidas a partir de programas o aplicativos informáticos, con base en el diseño de los cuadros y las especificaciones de procesamiento (a nivel de celda), de cada una de las variables o de los cruces de variables, desde la base de datos depurada de la operación estadística.

El proceso se desarrolla de la siguiente forma:

- ✓ El equipo temático de la encuesta envía las especificaciones para la generación de los cuadros de salida al equipo de sistemas.
- ✓ El grupo de sistemas programa estos resultados para la generación de cuadros en formato HTML o XLS.
- ✓ Revisados los cuadros generados por parte del equipo temático, se realizan ajustes a los programas, en caso de ser necesario.
- ✓ Se reciben los factores de expansión por parte del equipo de muestras y se generan los cuadros de salida definitivos, que permiten la visualización de los resultados de la encuesta.
- ✓ Los cuadros de salida se generan a nivel del total nacional, cabecera, centros poblados y rural disperso, y por departamentos.

2.5. DISEÑO DEL ANÁLISIS

Para el análisis y discusión de los resultados se tiene en cuenta el análisis estadístico y el análisis de contexto.

2.5.1. Métodos de análisis de resultados

- **Análisis de consistencia**

Este análisis incluye i) la comparación de resultados entre preguntas complementarias o relacionadas aplicadas en el mismo año, ii) el análisis del comportamiento de los principales indicadores básicos provenientes del módulo TIC de la Encuesta de Calidad de Vida ECV periodo 2023, para los dominios comunes (total nacional, cabecera, centros poblados y rural disperso) iii) y el contraste de los resultados por dominios (por ejemplo, cabecera frente a centros poblados y rural disperso, y departamentos entre sí).

Lo anterior permite identificar tendencias básicas, presencia de valores válidos, información atípica y consistencia entre variables. Estos análisis utilizan técnicas estadísticas de chequeo que permiten tener una visión más amplia y robusta del comportamiento del fenómeno estudiado.

- **Análisis de contexto**

El análisis de contexto consiste en la comparación de los resultados obtenidos en la ENTIC Hogares con los arrojados por otras investigaciones adelantadas el mismo año o un año cercano. La principal operación estadística que se usa en esa comparación es la Encuesta de Calidad de Vida (ECV) del DANE, la cual cuenta con preguntas comparables a las incluidas en la ENTIC Hogares y que, por su aplicación continua, facilitaría contar con un período de referencia similar.

- **Análisis de comparabilidad**

Con el propósito de contrastar la coherencia y la consistencia de los resultados y realimentar los análisis realizados, se define el uso de estándares estadísticos y se tienen en cuenta procedimientos de comparabilidad con otras operaciones estadísticas similares que se produzcan en el SEN, en otros países, o de organismos estadísticos internacionales (por ejemplo, de la OCDE y la UIT).

- **Análisis estadístico**

El análisis estadístico de los resultados de la ENTIC Hogares se efectúa en dos etapas:

Al inicio se realiza un análisis descriptivo a nivel de muestra. En ese análisis se examina la estructura de los indicadores por dominios de estudio a partir de la distribución de frecuencias y se detectan posibles inconsistencias y valores atípicos. Posteriormente, se verifica la cobertura en el nivel de segmento, viviendas, hogares y personas, y se observa la distribución de la pérdida de muestra para realizar los respectivos ajustes de la no respuesta.

Después se realiza el análisis inferencial para estimar los parámetros de la población objetivo a partir de la muestra. En este proceso se aplican los factores de expansión y se revisa la inferencia a la respectiva población objetivo, la cual es establecida para los correspondientes dominios de estudio. Además, se verifica que los ajustes del factor de expansión no generen sesgos en las estimaciones y que sus errores muestrales sean aceptables de acuerdo con los parámetros establecidos por dominios de estudio.

- **Análisis univariado**

En la ENTIC Hogares, se emplea principalmente distribuciones de frecuencias para las preguntas con varias opciones de respuesta, logrando presentarse como valores absolutos o en términos relativos, éstas suelen presentarse acompañadas de gráficos estadísticos (como histogramas, tortas, entre otras) para facilitar su comprensión y análisis.

- **Análisis bivariado**

Dentro del análisis estadístico de los datos, se tendrá en cuenta los límites (tanto inferior como superior) por medio de los cuales la estimación de los datos e indicadores puede oscilar, teniendo en cuenta, los coeficientes de variación de los datos calculados.

- **Análisis multivariado**

Teniendo en cuenta que la producción de información estadística en la ENTIC Hogares es en su mayoría descriptiva, no se empleará este tipo de análisis.

2.5.2. Anonimización de microdatos

El equipo de temática determina las variables que por confidencialidad deben ser anonimizadas en la base de datos que se va a publicar y a partir de la base de datos final, es decir, únicamente con las encuestas completas, se construye un archivo que es remitido a la Oficina de sistemas. Quien luego de realizar ese proceso, reenvía el archivo al área temática para su verificación, aprobación y posterior publicación en la página web del DANE.

La ENTIC Hogares no incluye en la base de datos algunas variables correspondientes a la identificación de las viviendas como son la dirección, barrio, número telefónico y datos de identificación de las personas tales como nombre y apellido, fecha de nacimiento, nombre de la empresa donde trabaja, entre otros.

En las bases compartidas por la Sala de Procesamiento Especializado Externo (SPEE), el protocolo de anonimización consiste en eliminar las variables de identificación de las fuentes (tales como

documento de identidad, barrio, nombre, dirección, teléfono, entre otras) y la información es publicada tal y como se recolectó, teniendo en cuenta que los usuarios firman acuerdos de confidencialidad en el manejo adecuado de la información estadística. Es fundamental destacar que la información disponible en la SPEE cumple con los protocolos de anonimización dispuestos por el DANE para garantizar la confidencialidad y uso de los datos sin violar la reserva estadística, de conformidad con el Código de Buenas Prácticas para las Estadísticas Oficiales. Además, cuenta con un equipo de profesionales encargados de verificar que los resultados generados por los investigadores preserven en todo momento la reserva estadística.

2.5.3. Verificación de la anonimización de microdatos

Para las bases que se comparten en la Sala de Procesamiento Especializado Externo (SPEE), el procedimiento consiste en verificar la eliminación de las variables de identificación de las fuentes (tales como documento de identidad, barrio, nombre, dirección, teléfono, entre otras). Adicionalmente, se incluye la documentación pertinente para que el usuario pueda consultar la metodología y diccionario de datos.

Para eliminar el riesgo de identificación de las fuentes, en este caso los hogares y personas que suministran los datos para la generación de la información estadística, se debe llevar a cabo un proceso de anonimización que permita proteger la privacidad de las fuentes y preservar el aprovechamiento de los datos a disposición de los usuarios. Por lo anterior, este procesamiento se basa en el documento Guía para la anonimización de bases de datos en el Sistema Estadístico Nacional publicado por el DANE.

2.5.4. Comités de expertos

Se realizarán reuniones a través de una mesa de trabajo de la ENTIC Hogares, conformada por personas expertas externas e internas del DANE, con el fin de realizar seguimiento a los principales indicadores sociales comparando con fuentes o estadísticas producidas por otras entidades. Esto permitirá hacer contraste entre las diferentes metodologías, sus coberturas y los resultados que cada una arroja. Este es un proceso pertinente puesto que así se garantiza la calidad de la información y la participación de los usuarios en el proceso de análisis y socialización de resultados.

Del mismo modo, se cuenta con una mesa de trabajo permanente con el MINTIC, como entidad líder del sector y secretaria técnica de la **Mesa Técnica de Estadísticas TIC**, en donde se encuentran los principales actores que participan en el seguimiento y desarrollo de la política de tecnología en el país.

2.6. DISEÑO DE LA DIFUSIÓN Y COMUNICACIÓN

Los resultados agregados de la ENTIC Hogares se publican en la página web del DANE, igualmente, el DANE cuenta con una Sala de Procesamiento Especializado Externo (SPEE), en la que se presta el servicio de acceso a los microdatos anonimizados de la operación estadística.

Para facilitar la correcta lectura e interpretación de los datos, los productos de difusión (Boletín técnico y Anexos), cuentan con una serie de gráficas y cuadros con sus respectivos comentarios y notas explicativas.

2.6.1. Diseño de sistemas de salida

El manejo y disposición de la información de la ENTIC Hogares está a cargo de la oficina de sistemas que cuenta con un sistema centralizado (servidor) para su almacenamiento y mantenimiento. Con el propósito de garantizar la seguridad de la información, este sistema cuenta un mecanismo de respaldo (o copia de seguridad) que permite la recuperación de la información en caso de ser necesario.

Este mecanismo consiste en realizar un proceso automático de backups del servidor donde reposa la base de datos usando tareas programadas del sistema operativo en ORACLE.

Almacenamiento en la Base de Datos Oracle: Una vez procesados, los datos en formato CSV se importan a la base de datos Oracle ENTIC 2024. Esta base de datos centraliza toda la información recolectada, proporcionando un repositorio único y seguro para el almacenamiento y gestión de los datos.

Análisis y Generación de Reportes: Los datos almacenados en la base de datos Oracle son analizados utilizando la herramienta SAS DANE. Este software de análisis estadístico permite realizar un análisis detallado de los datos, identificando tendencias y patrones significativos. Los resultados del análisis se utilizan para generar reportes detallados.

2.6.2. Diseño de productos de difusión y comunicación

Los productos e instrumentos de difusión de la ENTIC Hogares son:

- ✓ Boletín técnico.
- ✓ Anexos técnicos (Cuadros de Salida).
- ✓ Presentación de resultados en PowerPoint al comité de personas expertas convocado por las directivas del DANE.

Adicionalmente, como parte de la estrategia para difundir los resultados de la operación estadística, se realiza un comité externo donde se invitan a los principales usuarios externos de la información para presentar las principales cifras y análisis de las estadísticas a publicar.

Para los investigadores y usuarios en general, la base de datos anonimizada se encontrará disponible en la Sala de Procesamiento Especializado Externo (SPEE), previa activación de los protocolos que garantice la reserva estadística de la información.

2.6.3. Entrega de productos

Una vez finalizada la elaboración de los productos de publicación, se inicia el proceso de difusión y comunicación de resultados de la ENTIC Hogares, se establece la fecha de publicación y se divulga en el calendario de eventos de la página web del DANE, lugar donde se anuncian las fechas próximas de publicaciones.

El día de la publicación se publican en los siguientes sitios. Por un lado, en el link www.dane.gov.co, sección Estadísticas por tema, opción Tecnología e Innovación, enlace Encuesta de Tecnologías de la Información en Hogares (ENTIC Hogares), se accede a los productos que contienen resultados de la operación estadística.

2.6.4. Estrategia de servicio

Las solicitudes de información no disponible en los enlaces dispuestos para las consultas de los microdatos y los principales resultados de la ENTIC Hogares son recibidas de forma virtual (contacto@dane.gov.co), presencial (Banco de datos) o telefónica.

Por medio del Sistema de Gestión Documental, MERCURIO, esas solicitudes se asignan al equipo temático de Industria que debe responder a los requerimientos de manera oportuna y dentro de términos establecidos.

Para los usuarios que requieren un mayor nivel de información y tienen la capacidad de efectuar consultas expertas, el DANE pone a su disposición la sala de procesamiento especializado donde tendrá acceso a la base de datos anonimizada para hacer los cruces de variables que den respuesta a su solicitud.

2.7. DISEÑO DE LA EVALUACIÓN DE LAS FASES DEL PROCESO

De acuerdo con los controles de calidad y los informes de cobertura y seguimiento realizados durante el proceso de recolección de la ENTIC Hogares, se analizan y se determinan los puntos más importantes a tener en cuenta en los próximos ejercicios que se vayan a efectuar. Para esto se tiene en cuenta lo siguiente:

- La revisión de los informes presentados por cada una de las sedes y subsedes en donde se relacionan los problemas de recolección que se presentaron durante el operativo de campo, para así determinar los correctivos necesarios que se deben tomar en las futuras encuestas.
- La revisión de los informes de monitoreo o acompañamientos en campo realizados durante el desarrollo de la encuesta. Estos monitoreos permiten observar las dificultades metodológicas y conceptuales presentadas por el personal de recolección, así como los problemas de comprensión de preguntas por parte de las personas encuestadas, y así en el entrenamiento de las próximas encuestas, hacer énfasis en las preguntas que generaron dificultad y, si es necesario, profundizar en los diferentes manuales.
- Una vez consolidada la información, se lleva a cabo la revisión de frecuencias y la base de datos, lo cual permite determinar la necesidad en próximas encuestas de modificar o cambiar alguna opción de respuesta, que no haya sido muy clara o que haya generado algún inconveniente.

Durante cada una de las fases el proceso estadístico, la ENTIC Hogares recopila en actas o ayudas de memoria de las discusiones, observaciones y recomendaciones derivadas de las reuniones que se adelantan con expertos o grupos internos del trabajo, a partir de las cuales nacen las mejoras al proceso o sus herramientas. Igualmente, durante la etapa de recolección y validación de la información, se cuenta con herramientas en el aplicativo que permiten realizar un seguimiento por medio de los indicadores de calidad, cobertura y oportunidad descritos en los apartados anteriores.

Finalmente, los productos definitivos de publicación (boletín técnico y anexos) deben cumplir con la aprobación de los revisores de la Dirección Técnica, quienes evaluarán que sean productos de calidad con información que sea clara, consistente y acorde entre sí, para que cumplidos estos criterios sea aprobada su publicación.

2.8. DISEÑO DE LOS SISTEMAS DE PRODUCCIÓN Y FLUJOS DE TRABAJO

De acuerdo con el modelo funcional de la ENTIC Hogares, esta operación estadística tiene ocho (8) subprocesos, que se describen a continuación. En la descripción se identifican las funciones de cada uno de los equipos de trabajo involucrados, sus interrelaciones y los productos generados en la operación.

Detección y análisis de necesidades

Este proceso comienza con la recepción de los requerimientos de información provenientes de usuarios internos o externos frente a la ENTIC Hogares, los cuales llegan a la coordinación del GIT Temática Industria para ser analizados y evaluados en conjunto con el equipo temático. Una vez analizada la viabilidad técnico-económica para el desarrollo de la ENTIC Hogares, el GIT de Diseños muestrales define la muestra y con base en esta la Dirección de Recolección y Acopio procede a generar el presupuesto y cronograma de actividades para la operación estadística.

Diseño

Este subproceso comienza con el diseño temático a cargo del equipo de Tecnología e Innovación. El diseño temático comprende: la elaboración del formulario de recolección y del manual de conceptos básicos y recolección, a cargo del equipo temático, la elaboración del documento metodológico y la ficha metodológica, a cargo de todos los grupos de trabajo participantes, y la actualización de los demás manuales que son responsabilidad de la Dirección de Recolección y Acopio y Sistemas. En este subproceso el GIT de Diseños muestrales adelanta el diseño estadístico para la estimación y selección de la muestra. Así mismo y a partir del formulario final, el equipo temático define las normas de validación y consistencia para el diseño del aplicativo de recolección de información. Por último, todos los grupos participantes diseñan el programa de entrenamiento del personal que adelantará el operativo de campo.

Construcción

En este subproceso el GIT de Diseños muestrales procede a la selección de la muestra y al ensamble de las carpetas que contienen los segmentos seleccionados en los municipios, actividad realizada con el apoyo de la Dirección de Geoestadística, que proporciona la cartografía digital actualizada. Una vez finalizada esta actividad, la Dirección de Recolección y Acopio establece las rutas de recolección que seguirá el personal de campo en cada territorial a nivel nacional. Para finalizar, interviene la Oficina de Sistemas que se encarga de generar el aplicativo que permitirá la captura de la información, de acuerdo con las normas de validación establecidas por el equipo temático.

Recolección o acopio

Antes de comenzar la recolección de la información, desde el DANE Central se realiza el proceso de entrenamiento a los responsables de cada sede y subsede, de manera presencial o virtual, donde se imparten conocimientos temáticos, muestrales, de sistemas, logísticos y georreferenciación. Este entrenamiento es replicado en cada sede y subsede al personal de campo seleccionado. Antes de iniciar el operativo de campo se lleva a cabo el recuento de las edificaciones y las viviendas nuevas en la muestra, y una semana antes inicia el proceso de sensibilización a los hogares. En los hogares sensibilizados inicia la recolección de la información que cuenta con un riguroso proceso de supervisión con el fin de garantizar la calidad de la información. De acuerdo con los tiempos establecidos por la Dirección de Recolección y Acopio, las sedes y subsedes realizan el envío de la información a DANE Central, siguiendo los protocolos definidos por la Oficina de Sistemas. La información enviada es controlada por medio de los informes de cobertura reportados por cada sede y subsede.

Procesamiento

En este subproceso la Oficina de Sistemas realiza diariamente el cargue de la información enviada por las sedes desde el momento en que comienza la recolección, valida el cumplimiento de los criterios de completitud y hace revisión de consistencias. Así también, carga y comparte cada semana las tablas para que el equipo temático de la ENTIC Hogares realice su propio análisis de inconsistencias temáticas. Una vez finalizado el operativo y validada la información, se procede a consolidar la base de datos para que el GIT de Diseños Muestrales realice el cálculo y ajuste de los factores de expansión para generar los cuadros de salida.

Análisis

Una vez generados los cuadros de salida, estos son enviados por la Oficina de Sistemas al equipo temático de la ENTIC Hogares para realizar el análisis de coherencia. El paso siguiente lo realiza el GIT de Diseños Muestrales que se encarga de la generación de los errores de muestreo y significancia estadística de los datos que se van a publicar para que el equipo temático de la ENTIC Hogares produzca y analice la información. A partir de este punto y mediante un protocolo se anonimiza la base de datos. Al mismo tiempo se revisan y aprueban los boletines técnicos, junto con los anexos estadísticos (cuadros de salida) y la base anonimizada.

Difusión

La difusión de la información estadística está a cargo de la Dirección de Difusión y Cultura Estadística - DICE, que pone a disposición en la página web del DANE los productos estadísticos generados.

Evaluación

Este subproceso consiste en la realización de evaluaciones del desempeño al finalizar cada una de las fases del proceso estadístico explicadas anteriormente. En él deben participar todos los grupos involucrados en la operación estadística (Temático, Oficina de Sistemas, GIT de Diseños Muestrales, DIG, DICE y la Dirección de Recolección y Acopio).

3. DOCUMENTACIÓN RELACIONADA

Para la realización de la ENTIC Hogares se cuentan con la siguiente documentación:

- Metodología
- Ficha metodológica
- Diccionario de datos
- Formulario
- Especificaciones
- Plan de recolección
- Manual de usuario
- Manual de sistema
- Manual de diligenciamiento

GLOSARIO

Las principales definiciones que son presentadas a continuación fueron aprobadas en mesas de concertación con DIRPEN, área encargada del sistema de armonización de conceptos del DANE. Otras definiciones fueron adoptadas de referentes nacionales e internacionales.

- **Acceso a las TIC:** Comprende la infraestructura, los servicios, las aplicaciones, las condiciones y los medios para el acercamiento de grupos de interés a las tecnologías de la información y la comunicación (TIC). (Adaptado del glosario de conceptos del Ministerio de tecnologías de la información y la comunicación).
- **Banca electrónica y otros servicios financieros en línea:** Transacciones bancarias para pagos o transferencias, o para visualizar información sobre una cuenta bancaria. Incluye las transacciones electrónicas por Internet para otro tipo de servicios financieros como la compra de acciones, bonos y seguros. (Unión Internacional de Telecomunicaciones (2010). Definiciones de los indicadores mundiales de las telecomunicaciones/tic).
- **Brecha digital:** Comprende las diferencias en la oportunidad de acceso a las tecnologías de la información y la comunicación (TIC) y al uso de Internet entre individuos, hogares, empresas y

áreas geográficas de diversos niveles socioeconómicos. (OCDE (2006). Understanding the Digital Divide, OECD, 200, página 5).

- **Capacitación en TIC:** Se define la capacitación como el conjunto de procesos organizados, relativos tanto a la educación no formal o educación para el trabajo y el desarrollo humano (Ley 1064 de 2006) como a la informal, de acuerdo con lo establecido por la Ley General de Educación, dirigidos a prolongar y a complementar la educación inicial mediante la generación de conocimientos, el desarrollo de habilidades y el cambio de actitudes, con el fin de incrementar la capacidad individual y colectiva para contribuir al cumplimiento de la misión institucional, a la mejor prestación de servicios a la comunidad, al eficaz desempeño del cargo y al desarrollo personal integral. (Ministerio de Tecnologías de la Información y Comunicaciones de Colombia, 2024).
- **Chatbot:** Un chatbot es un programa informático que simula y procesa la conversación humana (ya sea escrita o hablada), permitiendo a los humanos interactuar con dispositivos digitales como si se estuvieran comunicando con una persona real. Los chatbots pueden ser tan simples como programas rudimentarios que responden a una consulta simple con una respuesta de una sola línea, o tan sofisticados como asistentes digitales que aprenden y evolucionan para ofrecer niveles cada vez mayores de personalización a medida que recopilan y procesan información” (Oracle, S.F).
- **Coefficiente de Variación Estimado (CVE):** Es una medida que resume la variabilidad de la estimación de un parámetro. Se obtiene a partir de la información de la muestra e indica el grado de precisión con el cual se está reportando un resultado. De tal forma que entre menor sea el CVE, menor incertidumbre se tiene de la estimación y ésta es más precisa. Es considerado una medida de dispersión relativa o de homogeneidad de los valores, y es de gran utilidad cuando se trata de comparar grupos de datos correspondientes a diferentes variables. (Departamento Administrativo Nacional de Estadística (DANE)).
- **Comercio electrónico:** Comprende la venta o la compra de bienes o servicios realizada mediante redes informáticas en plataformas electrónicas específicamente diseñadas con el propósito de registrar pedidos, independientemente de si el pago y la entrega de los bienes o servicios ocurren en línea. (Adaptado de OCDE. (2011) citado en el Consejo Nacional de Política Económica y Social CONPES No. 4012 (2020): Política Nacional de Comercio Electrónico. Recuperado el 20 de febrero del 2025).
- **Comprar u ordenar productos o servicios (actividad de uso de Internet):** Órdenes de compra colocadas mediante una plataforma electrónica a través de Internet, sin importar si el pago se realiza en línea o no. Incluye la compra por Internet de productos tales como música, viajes y alojamiento. No se incluyen las órdenes que se cancelaron o no fueron concretadas. (Departamento Administrativo Nacional de Estadística (DANE), 2021).
- **Consulta de medios de comunicación (actividad de uso de Internet):** Incluye el uso de radio o televisión vía Web y la búsqueda o consulta de revistas y periódicos en línea. (Departamento Administrativo Nacional de Estadística (DANE), 2021).

- **Correo y mensajería (actividad de uso de Internet):** Incluye el envío o recepción de correos electrónicos y llamadas telefónicas a través de protocolos de Internet. No se incluye la mensajería instantánea de las redes sociales. (Departamento Administrativo Nacional de Estadística (DANE), 2021).
- **Descargar software, imágenes, juegos, música o jugar en línea (actividad de uso de Internet):** Esta alternativa incluye el intercambio de juegos o juegos en línea, descargar música o imágenes, consultar horóscopos, descargar programas computacionales (software) y actualizaciones. No se incluye el uso de redes sociales ni la consulta de medios de comunicación. (Departamento Administrativo Nacional de Estadística (DANE), 2021).
- **Descargar o instalar programas computacionales (software):** Se refiere a poner en funcionamiento por primera vez dentro de un computador o tableta un programa computacional, como por ejemplo paquetes de ofimática (Word, Excel, Power Point, etc.), juegos, antivirus, navegadores de internet, sistemas operativos o aplicaciones. (Departamento Administrativo Nacional de Estadística (DANE), 2021).
- **Educación y aprendizaje (actividad de uso de Internet):** Esta alternativa incluye actividades formales de capacitación por internet, tales como estudios relacionados con cursos escolares o de educación terciaria, así como cursos de educación a distancia con actividades en línea. (Departamento Administrativo Nacional de Estadística (DANE), 2021).
- **Formación en habilidades y competencias digitales:** Se refiere al conjunto de procesos dirigidos a generar conocimientos, desarrollar habilidades, capacidades, destrezas y actitudes que se requieren para la utilización estratégica de la información, con el fin de ser capaces de trabajar, de relacionarnos con otras personas, con la administración, de comprar por Internet, de entretenernos, etc., todo ello con herramientas propias de las tecnologías digitales. (Ministerio de Tecnologías de la Información y Comunicaciones de Colombia, 2024).
- **Inteligencia Artificial:** Lenguaje de programación de computadoras avanzado que busca emular el razonamiento humano. (Adaptado de Glosario de términos estadísticos OCDE).
- **Internet:** Conjunto descentralizado de redes de comunicación interconectadas que utilizan la familia de protocolos TCP/IP para garantizar que las redes físicas heterogéneas que lo componen funcionen como una red lógica única de alcance mundial. (Concepto adaptado de las definiciones de ITU e INEGI, INEGI. Glosario ENTIC (2013)).
- **Lugar de uso:** Incluye el hogar, el trabajo, un establecimiento educativo, la casa de otra persona, un local con acceso comunitario a Internet, un local con acceso comercial a Internet, el desplazamiento de un sitio a otro y otros lugares. Las personas pueden responder respecto de más de un lugar. (Departamento Administrativo Nacional de Estadística, 2021).
- **Obtener información (actividad de uso de Internet):** Incluye la búsqueda y consulta de información sobre bienes o servicios, salud o servicios médicos y organizaciones gubernamentales en general. La información se puede obtener a través de los sitios Web o por correo electrónico. No incluye la consulta de medios de comunicación por Internet. (Departamento Administrativo Nacional de Estadística, 2021).

- **Redes sociales:** Es una comunidad de personas con un interés en común que utilizan un sitio Web u otras tecnologías para comunicarse entre sí y compartir información. Incluye mensajería instantánea (chats, blogs, foros, videos, etc.). (Departamento Administrativo Nacional de Estadística (DANE), 2021).
- **Suplantación de identidad (incidente digital):** La suplantación de identidad hace referencia al uso u apropiación indebida de la identidad de otro para actuar en su nombre. Esta suplantación se puede llevar a cabo mediante la utilización de tecnologías de la información y las comunicaciones tales como redes sociales, correos electrónicos, líneas telefónicas, mensajes de texto etc. Los objetivos del suplantador pueden ser múltiples, desde dañar la reputación, estafas, extorsiones o realizar ciberataques que afecten la infraestructura tecnológica de una organización u entidad. (Ministerio de Tecnologías de la Información y Comunicaciones de Colombia, 2024).
- **Tableta:** Dispositivo electrónico con características y habilidades similares a las de un computador portátil (como la navegación en Internet), pero con un menor rendimiento y un diseño más simple, compuesto básicamente por una pantalla táctil. (Adaptado de Instituto Nacional de Estadística y Geografía (INEGI). (s. f.). Glosario de definiciones).
- **Tecnologías de la Información y Comunicación (TIC):** Conjunto de recursos, herramientas, equipos, programas informáticos, aplicaciones, redes y medios que permiten la compilación, el procesamiento, el almacenamiento y la transmisión de información como voz, datos, texto, video e imágenes. (Ley 1341 de 2009, art. 6. m2009).
- **Teléfono celular convencional:** Dispositivo inalámbrico electrónico basado en la tecnología de ondas de radio (transmisión por radio frecuencia), y que tiene la misma funcionalidad que cualquier teléfono de línea fija, cuya principal característica es la realización de llamadas sin depender de ningún terminal fijo. (INEGI. (s.f.). INEGI.)
- **Teléfono celular inteligente:** Teléfono móvil celular construido sobre una plataforma informática y un sistema operativo móvil, y con características y habilidades similares a las de un computador, tales como almacenamiento de datos, conexión a Internet para correo electrónico y navegación, entre otras. (Adaptado de las definiciones de ITU e INEGI, INEGI. Glosario ENTIC (2013) INEGI. (s.f.). INEGI.)
- **Televisión, videos, películas u otro contenido audiovisual para entretenimiento (actividad de uso de Internet):** Incluye el uso de radio o televisión vía web únicamente con fines de entretenimiento y la búsqueda y consulta de películas o videos con el mismo fin. No se incluye el uso de redes sociales ni la consulta de medios de comunicación con fines de obtener información (Departamento Administrativo Nacional de Estadística, 2021).
- **Televisor:** Dispositivo que puede recibir señales de televisión utilizando medios comunes de acceso, por aire, cable o satélite. Excluye la función de TV integrada en otro dispositivo, como una computadora o un teléfono móvil. (Departamento Administrativo Nacional de Estadística, 2021).

- **Trámites con organismos gubernamentales (actividad de uso de Internet):** Incluye descargar o solicitar formularios, completar o presentar formularios en línea, hacer pagos en línea y compras por Internet a título de las organizaciones gubernamentales. No se incluye la consulta de información sobre las organizaciones gubernamentales. (Departamento Administrativo Nacional de Estadística (DANE), 2021).
- **Entrega de productos en línea en forma digitalizada:** Se refiere a bienes entregados a través de Internet únicamente en forma digitalizada, como por ejemplo libros digitales, informes, software, música, videos o juegos para computadora. A su vez, incluye también servicios entregados en línea, tales como servicios relacionados con la computación, servicios de información, reservas turísticas o servicios financieros. (Departamento Administrativo Nacional de Estadística (DANE), 2021).
- **Internet de las cosas: Internet of Things (IoT):** Dispositivos interconectados que recolectan e intercambian datos y pueden ser monitoreados o controlados remotamente mediante Internet, sin requerir la interacción entre humanos o entre humanos y computadores. Se excluyen los sensores y los dispositivos de detección simple de movimiento, sonido, temperatura, humo, etc. (Internet of things (IoT). Obtenido de TechTarget: (febrero de 2020)).
- **Informante idóneo:** Persona del hogar mayor de 18 años, que a falta del informante directo puede responder correctamente las preguntas. (Departamento Administrativo Nacional de Estadística (DANE), 2021).

BIBLIOGRAFÍA

Comisión Económica para América Latina y el Caribe (CEPAL). (2011). Caracterización del proceso de creación de estadísticas armonizadas sobre tecnologías de la información y la comunicación en los organismos nacionales de estadísticas.

Conferencia de las Naciones Unidas sobre Comercio y Desarrollo (UNCTAD). (2009). *Manual para la producción de estadísticas sobre la economía de la información*. Naciones Unidas.

Departamento Administrativo Nacional de Estadística. Sistema de Consulta de Conceptos Estandarizados.

Departamento Administrativo Nacional de Estadística. (2023). Encuesta de Calidad de Vida (ECV).

Departamento Administrativo Nacional de Estadística. (2024). Encuesta de Calidad de Vida (ECV).

Departamento Administrativo Nacional de Estadística. (2021). Encuesta TIC en empresas ENTIC.

Hansen, et al. (1953). *Sample survey methods and theory*. Vol. 1. Methods and applications.

Instituto de Estadística de la Unesco (UNESCO Institute for Statistics). (2009). Medición de las tecnologías de la información y la comunicación (TIC) en educación: Manual del usuario.

Observatorio para la Sociedad de la Información en Latinoamérica y el Caribe (OSILAC). (2004). El estado de las estadísticas sobre Sociedad de la Información en los Institutos Nacionales de Estadística de América Latina y el Caribe.

Organización de las Naciones Unidas. (2005). Indicadores clave de las tecnologías de la información y de las comunicaciones.

Organización de las Naciones Unidas. (2010). Indicadores clave de las tecnologías de la información y de las comunicaciones.

Organización para la Cooperación y el Desarrollo Económicos (OCDE). (2006). Understanding the Digital Divide.

Organización para la Cooperación y el Desarrollo Económicos (OCDE). (2015). Perspectivas de la OCDE sobre la economía digital.

http://www.oecd.org/sti/ieconomy/DigitalEconomyOutlook2015_SP_WEB.pdf

Organización para la Cooperación y el Desarrollo Económicos (OCDE). (2019), Cómo medir la transformación digital: Hoja de ruta para el futuro, OECD Publishing, Paris/ACUI, Barranquilla, <https://doi.org/10.1787/af309cb9-es>.

Organización para la Cooperación y el Desarrollo Económicos (OCDE). (2019), Going Digital: Shaping Policies, Improving Lives, OECD Publishing, Paris, <https://doi.org/10.1787/9789264312012-en>.

Organización para la Cooperación y el Desarrollo Económicos (OCDE). (s.f.). Observatorio de Inteligencia Artificial.

Oracle. (s.f.). What is a Chatbot?

Partnership on Measuring ICT for Development - ITU. (2010). Indicadores clave sobre TIC.

Sánchez, et al. (2012). La Sociedad de la Información: Génesis, Iniciativas, Concepto y su Relación con las TIC. Revista UIS Ingenierías.

Unión Internacional de Telecomunicaciones (UIT). (2010). Definiciones de los Indicadores Mundiales de las Telecomunicaciones.

Unión Internacional de Telecomunicaciones (UIT). (2014). Manual para la medición del uso y el acceso a las TIC por los hogares y las personas.

Unión Internacional de Telecomunicaciones (UIT) (2020). Manual para la medición del uso y el acceso a las TIC por los hogares y las personas. Ginebra – Suiza.

ANEXOS
A. Cuadros de salida según principales indicadores esperados de la ENTIC HOGARES:

ENCUESTA DE TECNOLOGÍAS DE LA INFORMACIÓN Y LAS COMUNICACIONES EN HOGARES - ENTIC HOGARES 2024	
Tenencia de bienes y servicios TIC en hogares	
1 C.1	Tenencia de bienes y servicios TIC en los hogares. Total nacional, áreas y departamentos
2 Mod.2-1	Conocimiento de establecimientos para la venta de equipos TIC. Total nacional, áreas y departamentos
3 Mod.2-2	Características de la conexión a Internet en el hogar. Total nacional, áreas y departamentos.
4 Mod.2-3	Uso y servicios para ver televisión en los hogares. Total nacional, área y departamentos.
5 Mod.2-4	Razón principal por la que el hogar no tiene conexión a Internet. Total nacional, área y departamentos.
6 Mod.2-5	Razón principal por la que el hogar no tiene computador (de escritorio, portátil o tableta). Total nacional, área y departamentos
Uso de bienes y servicios TIC por personas	
7 C.4	Uso del computador, tableta, Internet, teléfono celular y señal de radio dentro del hogar, según departamentos del país y sexo. Personas de 5 y más años de edad
8 Mod.4-1	Frecuencia de uso de computador de escritorio, portátil, tableta, teléfono celular y escucha de radio dentro del hogar, según departamentos del país y sexo.
9 Mod.4-2	Actividades que saben realizar las personas en un computador* o cuando usa Internet, según departamentos del país y sexo.
10 Mod.4-3	Frecuencia de uso de Internet (desde cualquier lugar y dispositivo) y promedio de horas de uso al día, por sexo. Total nacional, área y departamentos.
11 Mod.4-4	Uso de dispositivos para acceder a Internet, por sexo. Total nacional, área y departamentos.
12 Mod.4-5	Uso de dispositivos para acceder a Internet según principal finalidad y tiempo que lo utiliza, por sexo. Total nacional, área y departamentos.
13 Mod.4-6	Sitios de uso de Internet y percepción, por sexo. Total nacional, área y departamentos.
14 Mod.4-7	Actividades de uso de Internet, por sexo. Total nacional, área y departamentos.
15 Mod.4-8	Actividades en redes sociales, por sexo. Total nacional, área y departamentos
16 Mod.4-9	Actividades de uso de servicios en la nube, por sexo. Total nacional, área y departamentos.
17 Mod.4-10	Actividades de uso de herramientas de Inteligencia Artificial a través de Internet y principal razón de uso por sexo. Total nacional, área y departamentos.
18 Mod.4-11	Principal razón de no uso de herramientas de Inteligencia Artificial a través de Internet, por sexo. Total nacional, área y departamentos.
19 Mod.4-12	Trámites, servicios o consultas con entidades públicas que las personas realizaron a través de Internet, por sexo. Total nacional, área y departamentos.
20 Mod.4-13	Percepción de las personas sobre los trámites en línea con entidades públicas, por sexo. Total nacional, área y departamentos.
21 Mod.4-13A	Frecuencia de uso de sitios web del Estado, por sexo. Total nacional, área y departamentos.
22 Mod.4-14	Conocimiento y participación de programas de formación en habilidades y competencias digitales, por sexo. Total nacional, área y departamentos.
23 Mod.4-15	Razones por las cuales las personas no realizan trámites en línea con entidades públicas, por sexo. Total nacional, área y departamentos.
24 Mod.4-16	Bienes y servicios comprados a través de Internet según tipo de bien, medio de pago, método de entrega y calificación del proceso de compra, por sexo. Total nacional, área y departamentos.
25 Mod.4-17	Razones por las cuales no se realiza compras de productos o servicios a través de Internet, por sexo. Total nacional, área y departamentos.
26 Mod.4-18	Razones por las cuales la persona no utiliza Internet, por sexo. Total nacional, área y departamentos.
27 Mod.4-19	Tenencia y uso de teléfono celular según medios por el que accede, frecuencia y actividades de uso, por sexo. Total nacional, área y departamentos.
28 Mod.4-20	Escucha de señal de radio dentro del hogar según frecuencia y actividad, por sexo. Total nacional, área y departamentos.
29 Mod.4-21	Conocimiento y uso de código postal, por sexo. Total nacional, área y departamentos.
Seguridad y uso responsable del Internet	
30 Mod.5-1	Hábitos y prácticas de protección de información en Internet, por sexo. Total nacional, área y departamentos.
31 Mod.5-2	Prácticas de control parental en Internet, por sexo. Total nacional, área y departamentos.
32 Mod.5-3	Indicentes de seguridad digitales, por sexo. Total nacional, área y departamentos.

**METODOLOGÍA GENERAL****CÓDIGO: DSO-ENTIC_H-MET-001**
VERSIÓN: 1

PROCESO: Producción Estadística

OPERACIÓN ESTADÍSTICA: Encuesta de Tecnologías de la Información y las Comunicaciones en Hogares- ENTIC_H

VERSIÓN	FECHA	RAZÓN DE LA ACTUALIZACIÓN
1	28/Julio/2025	Creación de documento.

ELABORÓ	REVISÓ	APROBÓ
Nombre: Edgar Eduardo Guayazan Sierra Cargo: Profesional Especializado	Nombre: Zaura Sierra Hernández Cargo: Coordinadora GIT Temática de Industria	Nombre: Cesar Mauricio López Alfonso Cargo: Director Técnico

Si este documento es impreso se considera copia no controlada