



MODELO DE GOBERNANZA DE LA **INFRAESTRUCTURA DE DATOS** PARA EL DESARROLLO DE TECNOLOGÍAS EMERGENTES



El futuro
es de todos

Gobierno
de Colombia

MODELO DE GOBERNANZA DE LA INFRAESTRUCTURA DE **DATOS** PARA EL DESARROLLO DE TECNOLOGÍAS EMERGENTES

Documento borrador para discusión

Autor:

Armando Guío Español

Consultor de la Corporación Andina de Fomento

Supervisores:

Víctor Manuel Muñoz Rodríguez

Director Departamento Administrativo de la Presidencia

Carlos Santiso

Director de Innovación Digital del Estado en CAF

María Isabel Mejía Jaramillo

Ejecutiva Senior en Gobierno Digital, Innovación Digital del Estado en CAF

Fecha:

Abril, 2021

Entidades involucradas en la realización del documento



El futuro
es de todos

Gobierno
de Colombia



El futuro
es de todos

Consejería Presidencial
para asuntos económicos
y transformación digital

CONTENIDO

+

I INTRODUCCIÓN

pág 4-6

II MODELO DE GOBERNANZA DE DATOS EN EL SECTOR PÚBLICO

pág 7-17

III UN NUEVO CONTEXTO Y OBJETIVOS DE POLÍTICA PÚBLICA PARA LA GOBERNANZA DE DATOS

pág 18-20

IV HACIA UN NUEVO MODELO DE GOBERNANZA DE DATOS EN COLOMBIA

pág 21-26

V GENERANDO CONFIANZA EN LA CREACIÓN DE UNA INFRAESTRUCTURA DE DATOS

pág 27-29

VI NECESIDAD DE INTERCAMBIAR MAYOR INFORMACIÓN CON EL SECTOR PRIVADO Y TENER SUS APORTES EN LA CONSTRUCCIÓN DE LA INFRAESTRUCTURA

pág 30-32

VII PROPUESTA DE GOBERNANZA PARA LA INFRAESTRUCTURA DE DATOS PÚBLICOS EN COLOMBIA

pág 33-46



La necesidad de crear una infraestructura de datos robusta parte de la importancia que tiene este insumo en la generación de tecnologías emergentes, como la inteligencia artificial. Hoy en día los datos empiezan a convertirse en un activo estratégico de cualquier país y Colombia ha venido realizando esfuerzos importantes por su desarrollo y para facilitar su acceso. Dada esta importancia transversal, es necesario proponer un marco específico para la creación de una infraestructura que cumpla las siguientes características: que sea robusta, que facilite la interoperabilidad, útil para el desarrollo y diseño de las tecnologías emergentes y con el propósito de servir al bien común y al desarrollo sostenible del país.

Para este fin se busca que el conjunto de datos públicos tenga gobernanza que sea capaz de generar una infraestructura sostenible y que tome decisiones estratégicas para su mantenimiento. Es claro que el país ha realizado importantes avances para la generación de una infraestructura de datos, especialmente en el sector público a través de varias iniciativas y políticas, muchas derivadas del documento Conpes 3920 de 2018, también conocida como la Política Nacional de Explotación de Datos (*Big Data*). Esta misma política hizo énfasis en la necesidad de desarrollar una institucionalidad que respondiera al desarrollo de dicha infraestructura.

Ahora bien, importantes cambios se han presentado desde la expedición del documento Conpes de *Big Data*. En primer lugar, el Estado Colombiano ha vivido transformaciones en el campo de la política pública de tecnología e innovación. Los cambios institucionales generan nuevos actores que tienen una agenda y preocupación particular y que empiezan a interactuar con entidades que tradicionalmente han estado encargadas de estos temas. Por ejemplo, hoy en día Colombia tiene una Consejería Presidencial exclusivamente dedicada a la Transformación Digital y un Ministerio de Ciencia, Tecnología e Innovación, algo que no se preveía en su momento dentro de la redacción del documento Conpes 3920.

Por otro lado, el documento Conpes 3920 buscaba desarrollar una infraestructura de datos completa y robusta que ante todo mejorara los sistemas de datos abiertos y permitiera desarrollar sistemas de *big data* dentro de las entidades públicas. Sin embargo, hoy en día dicha infraestructura debe dirigirse a cumplir con las necesidades de mercados y tecnologías emergentes, especialmente bajo un contexto no solo de digitalización o *datafificación*, sino de cuarta revolución industrial. Así mismo, esta infraestructura tiene ahora como principal objetivo la transformación digital del Estado y la implementación de diversos proyectos transformacionales que se han generado al respecto y que se valen de las últimas innovaciones disponibles. Así mismo, en la actualidad Colombia tiene un plan de transformación digital e inteligencia artificial materializada en otro documento Conpes (3970). Bajo este contexto, la infraestructura de datos sufre una serie de transformaciones en su propósito, objetivos y en el marco institucional aplicable.

La toma de decisiones frente a la forma como debe organizarse esta infraestructura datos también debe cambiar bajo este nuevo contexto. Es por esta misma razón que el mismo documento Conpes 3975 de 2019 señaló la necesidad de dirigir los esfuerzos de implementación del documento Conpes de *Big Data* para que responda a las necesidades del cumplimiento de las metas principales del plan de inteligencia artificial que el país ha desarrollado para los próximos años:

“En décimo octavo lugar, el Ministerio de Tecnologías de la Información y las Comunicaciones en conjunto con la Presidencia de la República implementará todas las medidas destinadas a la creación de la infraestructura de datos como resultado de la ejecución del Documento CONPES 3920 Política Nacional de Explotación de Datos. En este trabajo se hará especial énfasis en la implementación de los lineamientos para la explotación de datos en el sector público y el mercado de datos.”

Así mismo, se presenta una preocupación de ahondar en mecanismos que permitan que la infraestructura de datos del país también pueda nutrirse cada vez más de los datos recolectados y organizados por el sector privado. Se debe profundizar en mecanismos de colaboración que permitan una mayor interacción de este sector y que estos sean considerados en la toma de decisiones y en los estándares y criterios utilizados para utilizar esta información. Si Colombia tiene el objetivo de crear un mercado de inteligencia artificial, es necesario que la gobernanza de los datos obedezca también a esas necesidades, especialmente de un sector emprendedor que quiere ser competitivo mediante el acceso a este recurso.

Por último, el contexto actual ha demostrado la necesidad de recolectar y procesar una gran cantidad de datos dentro del sector público, incluidos datos de carácter personal. En este escenario la generación de confianza es esencial, en especial cuando estos datos tienden a ser públicos. La gobernanza de datos puede llevar a ser un elemento que genere mayor confianza al respecto. Al final lo que se espera es que un sistema de gobernanza de datos públicos también implemente un sistema de pesos y contrapesos que permita que la formación de una infraestructura de datos que también involucra datos personales sea generadora de confianza y brinde posibilidades de participación y control por parte de la ciudadanía.

Todos estos son elementos que hacen necesaria la definición de un modelo de gobernanza específico para el desarrollo de una infraestructura de datos, que esté especialmente dirigida al desarrollo de sistemas de inteligencia artificial y otras tecnologías emergentes que se beneficien de este recurso.

+

III

MODELO DE GOBERNANZA DE DATOS EN EL SECTOR PÚBLICO



Es claro que un modelo de gobernanza que beneficie e impulse el uso de tecnologías emergentes como la inteligencia artificial, tiene distintas capas y no se compone solo de decisiones técnicas. Como lo ha señalado la OECD la deficiencia de muchos modelos de gobernanza de datos desarrollados por distintos Estados se presentan por estar estrictamente dirigidos a obedecer criterios técnicos. La OECD concluye que un enfoque sólido en aspectos técnicos como el resultado primario de la gobernanza de datos puede orientar mal las decisiones de política relacionadas con datos (OECD, 2019). Por ejemplo, al enfocarse principalmente en la adopción de soluciones tecnológicas como interfaces de programación de aplicaciones (API) y estándares de datos, en vez de además habilitar el contexto organizacional adecuado de gobernanza y contexto cultural para hacer que estas herramientas sean valiosas para dirigirse a cambios de política (OECD, 2019). Todos estos elementos son clave para una buena gobernanza de datos (OECD, 2019).

A partir de lo anterior, se puede afirmar que un modelo de gobernanza debe ser comprensivo y holístico, buscando abordar el uso de los datos desde distintos ángulos y sin el objetivo de generar una única forma de organización de la información que sea rígida y poco adaptable a nuevas necesidades. De esta forma, lo que se busca es generar una organización y coordinación estatal capaz de ser flexible y que se adapte a las necesidades que pueden ir surgiendo alrededor del uso de la información. Así mismo, es necesario que el modelo de gobernanza permita la toma de decisiones de manera informada y considerando las distintas implicaciones que puede tener una decisión que afecte la formación de una infraestructura de datos.

Ahora bien, bajo el escenario de una cuarta revolución industrial, el Gobierno colombiano debe evitar la posibilidad de caer en modelos centralizados y poco adaptables a los cambios que se puedan presentar. En momentos la creación de modelos de gobernanza tiende a crear la idea de que los países deben centralizar su información y generar estándares uniformes para este fin.

Por lo tanto, este documento partirá por definir lo que un modelo tradicional de gobernanza de datos contiene y lo que lo caracteriza. Una vez presentadas las características propias de estos modelos se presentará la forma como se desea hacer una transición en Colombia hacia un modelo ontológico novedoso que permita desarrollar una infraestructura de datos novedosa y especialmente centrada en la interoperabilidad legal y regulatoria, no solo de los datos que componen dicha infraestructura.

1 | Características generales de un modelo de gobernanza de datos y casos de estudio

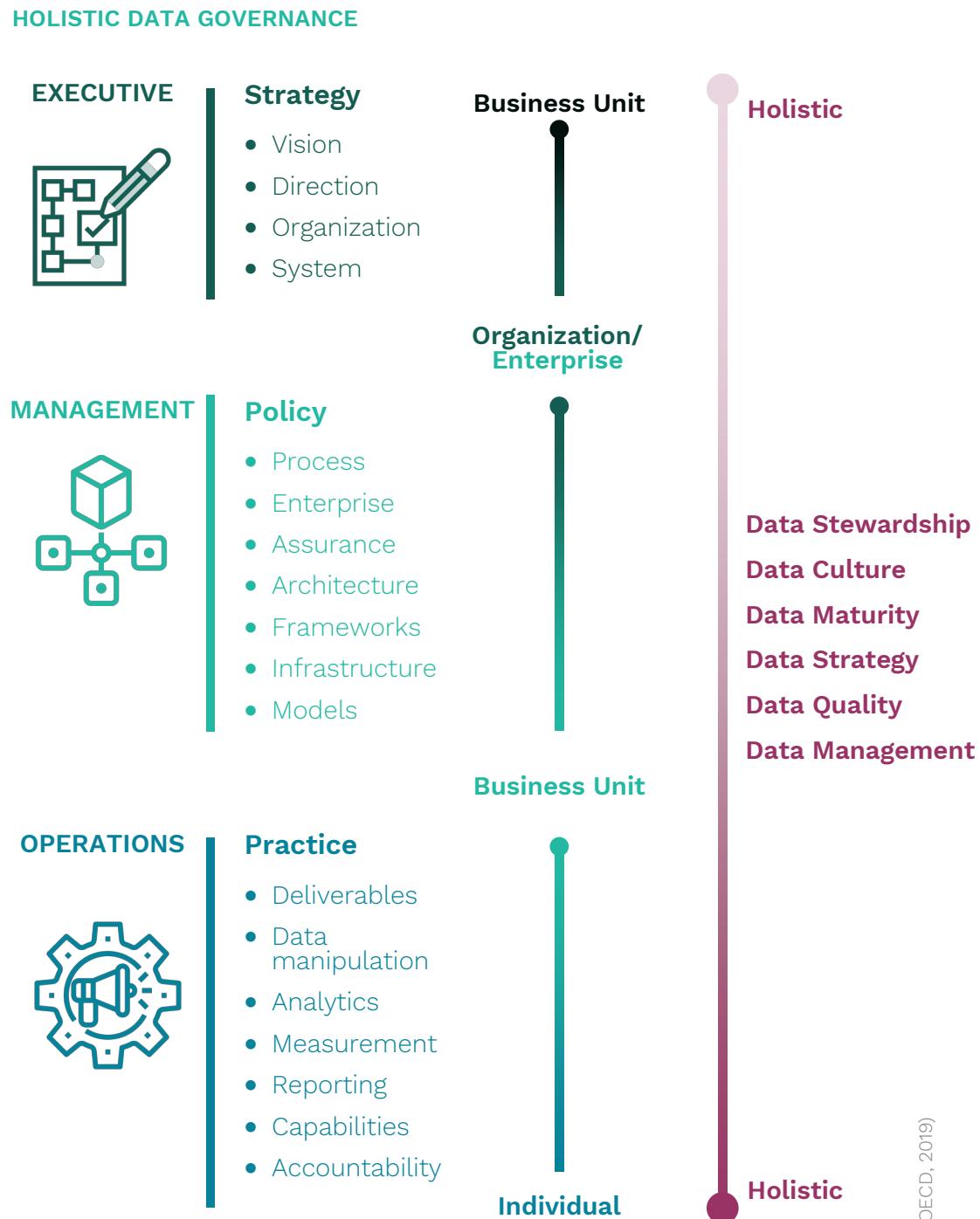
Un modelo de gobernanza de datos está íntimamente ligado con el ciclo de vida de los datos. Como lo señala el mismo término, el ciclo de vida busca definir la forma como los datos se crean, recolectan, circulan, procesan y son eliminados dentro de una organización. A partir de este proceso es posible determinar las áreas o sujetos involucrados en dichos procesos y la forma como realizan su trabajo. De esta manera, es posible establecer quiénes toman decisiones frente a la información, qué directrices o políticas han desarrollado o siguen en el cumplimiento de sus funciones y cómo deberían interactuar entre sí. De ahí la importancia de conocer las características propias de la infraestructura de datos y de los ciclos de vida que caracterizan a la información que la componen.

Ahora bien, generalmente la gobernanza de datos que proponen los países busca ir más allá, señalando que las decisiones frente a los datos en distintas organizaciones no recaen exclusivamente en temas técnicos, sino que dependen incluso de la alta gerencia de una entidad. Esto incluso hace que se genere una nueva cultura dentro de las organizaciones públicas que permita entender la importancia en la toma de decisiones frente a la forma como se debe administrar un recurso estratégico.

Esto significa que los operadores que interactúan constantemente con la información no son los únicos responsables de su uso y procesamiento, sino que al final están siguiendo las directrices, políticas e incluso estándares éticos definidos por la alta dirección. De esta forma, se busca que los órganos directivos asuman la importancia que tienen en el uso de este activo y definan propuestas generales sobre la forma como debe utilizarse.

Es así como se generan modelos holísticos de la gobernanza de datos que no solo se centran en la operación técnica de la información, sino en la forma como las decisiones estratégicas y de dirección de una organización interactúan con las disposiciones técnicas. En el caso del Estado, a esto se une el marco legal, normativo y regulatorio que hace posible la existencia de una infraestructura de datos y su gobernanza específica.

Ejemplo de esto es el modelo propuesto por el Gobierno de Nueva Zelanda y que resalta las diferencias en las distintas etapas de gobernanza, siendo uno de los más claros y comprensibles a nivel mundial:



Como se puede apreciar, el gobierno neozelandés resalta tres actividades principales que caracterizan su modelo de gobernanza de datos: (i) decisión ejecutiva; (ii) administración y gestión de la información; y (iii) operaciones que involucran estos datos. De cada una de estas actividades se derivan distintas tareas específicas como la creación de estrategias, políticas o modelos y prácticas ha ser desarrolladas e implementadas por los funcionarios que interactúan con la información. Se puede observar que este es un modelo basado en niveles de abstracción, yendo de una visión mas general a la implementación de la misma a través de modelos de arquitectura y prácticas más específicas, incluso con responsabilidades específicas en una persona o individuo. Este modelo permite evidenciar la forma como la distribución general de tareas dentro del modelo de gobernanza hace que las tareas más específicas respondan a objetivos generales. Esto además demuestra la forma como se conecta la visión u organización de la infraestructura de datos, lo cual tendrá un impacto directo en los entregables que se generan o los estándares y medidas de procesamiento de los datos y de reporte de las actividades que involucren a esta información. De ahí que los documentos de política pública y de principios generales tengan un alto impacto en los procedimientos técnicos y en los estándares que se utilizan para organizar una infraestructura de datos (Sweeney, 2020).

Es así que se han generado distintos modelos de gobernanza a nivel mundial que en gran medida siguen estas características. Es por esto que en el año 2017, la British Academy y la Royal Society del Reino Unido publicaron un informe conjunto en el que definieron las principales categorías y funciones que caracterizan a un modelo de gobernanza de datos, especialmente para aquellos que responden a sistemas que se están gestando de forma rápida y cambiante (ej, ambientes digitales). Dentro de estas vale la pena resaltar las siguientes:

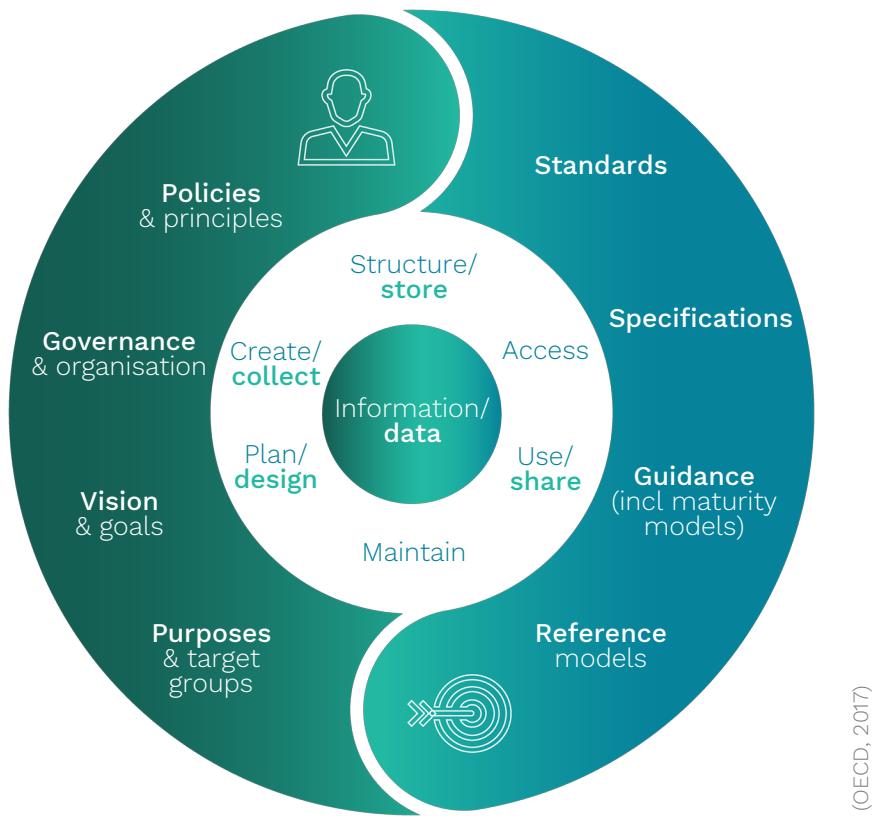
- **Anticipar, monitorear y evaluar:** considerando futuros alternativos, gestionando riesgos, manteniendo la calma con los cambios y reflexionando sobre el desempeño (The British Academy and the Royal Society, 2017).
- **Construir prácticas y establecer estándares:** habilitando y mejorando continuamente las prácticas adecuadamente fundamentadas que pueden extenderse rápidamente a través de sectores y usos relevantes (The British Academy and the Royal Society, 2017).
- **Aclarar, aplicar y subsanar:** asegurando los arreglos suficientes para recolectar evidencia, debatir y tomar decisiones y para acciones como incentivos, permisos, remedios para daños y penalidades. La aplicación adecuada de los Principios para la Gobernanza de Datos a estas funciones deja claro que involucrar a los *stakeholders* más significativos es esencial para desarrollar y llevar a cabo todas estas funciones (The British Academy and the Royal Society, 2017).

Eso incluso lleva a la British Academy y la Royal Society a establecer las distintas entidades del sector público y de la sociedad civil que están relacionadas con estas tareas dentro del Reino Unido, mostrando la complejidad y diversidad que caracteriza tradicionalmente a estos sistemas de gobernanza. Los actores que llevan a cabo estas funciones, que son públicos, privados y de la sociedad civil, incluyen la Information Commissioner's Office (ICO), la UK Statistics Authority, agencias de fondos para investigaciones, organizaciones no gubernamentales (ONG), universidades, la rama judicial, órganos de industria y asociaciones profesionales (The British Academy and the Royal Society, 2017). Algunos actores son sectoriales, como el rol del National Data Guardian con respecto a los datos en usos médicos, mientras que otros como la ICO trabajan de manera transversal a los sectores (The British Academy and the Royal Society, 2017). La gobernanza en el Reino Unido también se encuentra informada por estándares internacionales (The British Academy and the Royal Society, 2017).

Estos elementos se han materializado en modelos de gobernanza de datos propuestos en varios países. Por ejemplo, Estonia, que a través X-Road, y en la actualidad con X-Tee, viene generando una serie de pasos hacia un modelo interoperable sostenido por una gobernanza que se enfoca en tres puntos esenciales: fuentes de información, procesamiento y almacenamiento de información y propósitos definidos (e-estonia, 2018; Republic of Estonia Information System Authority, 2019).

El modelo resalta la importancia de identificar y vincular diferentes fuentes de datos (e.g. datos del sector privado, datos administrativos y datos del censo) a diferentes tipos de uso de datos (e.g. análisis de policía, investigación, operacional), para asegurar de manera estratégica el manejo y almacenamiento de datos adecuado (OECD, 2019). Hay cuatro retos principales que se identifican como cruciales para enfrentar la creación de un mejor marco de gobernanza de datos: la recolección, el almacenamiento, el crecimiento y la entrega de los datos (OECD, 2019). Estos retos cubren una sección grande de la cadena de valor de los datos, desde entender los activos de datos y establecer los principios de gobernanza de datos para el tratamiento, transmisión y diseminación de meta datos y datos (OECD, 2019).

Por su lado, Noruega ha centrado su modelo en los datos y en el ciclo de vida de estos, para luego establecer la forma como las políticas, estándares y prácticas rodean los propósitos del ciclo de vida descrito:



(OECD, 2017)

Como parte de su trabajo en el desarrollo de la arquitectura nacional de TI en Noruega, la Agencia para el Manejo Público y Gobierno Electrónico creó un modelo de información de gobernanza que posicionó el manejo de datos del sector público como el centro de la transformación digital del sector público noruego (OECD, 2017). Al posicionar los datos en el corazón del modelo de gobernanza de información y para complementar con visiones estratégicas, políticas, principios, estándares y guía para el mejor uso de los datos del sector público, los órganos públicos en Noruega cuentan con herramientas que permiten apalancar los datos como un activo estratégico para la toma de decisiones (OECD, 2017).

A partir de los varios modelos presentados y analizados por entidades como la OCDE, es posible apreciar que lo que se busca ante todo es generar estándares que permitan formar una infraestructura pública de datos de fácil acceso y que se dirijan a un modelo de estandarización. Estos modelos tienden a ser holísticos rodeando la parte técnica de una serie de decisiones de política pública que guían y acompañen este proceso. Esto se hace especialmente relevante cuando se aprecia la forma como esta infraestructura incluso puede verse construida con datos personales, frente a los cuales surgen varias inquietudes de índole legal, teniendo en cuenta los derechos fundamentales que interactúan.

Así mismo, las nuevas tecnologías que han surgido y que han tenido un amplio despliegue en la última década, como el *cloud computing*, generan nuevas inquietudes frente a la forma como se alberga información pública en servidores privados y las distintas condiciones que deben darse para facilitar este proceso. A esto responden también la necesidad de generar sistemas interoperables que maximicen la eficiencia de la infraestructura generada y faciliten los procedimientos de acceso y uso. Ahora bien, los Gobiernos cada vez entienden más que este no es un fenómeno propio y exclusivamente técnico, sino que también incluye una serie de decisiones políticas y de gobierno.

Como lo señala Urs Gasser, profesor de la Universidad de Harvard, los gobiernos pueden formar una parte importante del nivel institucional de interoperabilidad en su rol como cuidadores de un ambiente legal robusto y estable (Gasser, 2015). El éxito futuro de sistemas complejos emergentes, como el *cloud computing*, dependerá no solo de las fuerzas del mercado, sino también de un marco legal bien desarrollado (Gasser, 2015). Este marco legal debe establecer confianza y seguridad jurídica tanto para usuarios como para prestadores de sistemas interoperables futuros (Gasser, 2015). Las relaciones entre la interoperabilidad y la ley son muchas, complejas y enredadas (Gasser, 2015). La ley puede permitir establecer, ajustar o mantener la interoperabilidad; al mismo tiempo, la interoperabilidad es también una característica del sistema legal mismo, entendido como interoperabilidad legal (Gasser, 2015). La interoperabilidad legal, definida en términos amplios, es el proceso de hacer que las normas legales funcionen en conjunto a través de jurisdicciones (Gasser, 2015). Esta interoperabilidad puede ocurrir dentro del sistema legal de una sola nación o a través de fronteras nacionales (Gasser, 2015). Como la interoperabilidad técnica, la interoperabilidad legal no es una meta en sí misma, sino un medio para uno o más fines (Gasser, 2015).

Como se verá más adelante estos son muchos de los elementos que han caracterizado las decisiones de gobernanza de datos y de desarrollo de infraestructura pública de datos en Colombia, sin que todavía exista un modelo definido de cooperación entre entidades. Así mismo, muchos de estos elementos han llevado a los países a considerar la necesidad de generar estrategias de protección de su infraestructura. Ejemplo de esto son las medidas de localización de datos que los países justifican de distintas formas: protección de la privacidad de sus ciudadanos, la competitividad de sus emprendedores digitales o la seguridad nacional, entre otras razones. Estas medidas han llevado a generar limitaciones a la salida de alguna de la información de un territorio específico. Todo estos son elementos y discusiones de política pública que caracterizarán muchos de los modelos de gobernanza que se generan alrededor del mundo.

2 | La aproximación a la gobernanza de datos en Colombia

Durante la última década se han venido realizando esfuerzos importantes para construir una infraestructura de datos, siendo uno de los principales esfuerzos la expedición en el año 2018 del documento Conpes 3920 de *Big Data*. Este fue uno de los primeros documentos públicos en Colombia en hacer referencia a la Gobernanza de datos y sus características. Este documento obedece a un esfuerzo por lograr una cultura de datos o la *datafificación* del sector público.

Dentro de la política se definió la gobernanza de datos, en el contexto de una política pública de digitalización del Estado, de la siguiente manera:

“Es el conjunto de reglas que gobiernan el ciclo de vida y flujo de los datos de acuerdo con su tipología (World Economic Forum, 2015) con el fin de garantizar su calidad, uso, titularidad, compartición, seguridad y supresión, y que están orientadas a la generación de valor, minimizando los riesgos asociados a ello. Esta se expresa a través de una política que articula los controles, las metas de negocio, objetivos estratégicos, los procesos involucrados e indicadores de gestión. La gobernanza enmarca los procesos de generación, recolección, compartición, agregación, explotación e innovación, que se describen a continuación.”

Varios elementos se extraen de la definición propuesta en este momento:

- La gobernanza está intrínsecamente relacionada con el **ciclo de vida de los datos** y responde a las necesidades y procesos que rodean dicho ciclo.
- La gobernanza se entiende como la forma de **expresar y materializar una política de datos**. Es por esto por lo que el documento Conpes 3920 casi que podría ser considerado como el marco de gobernanza de datos públicos, algo que como ya se ha visto puede ser altamente ambicioso, pero deja de lado una perspectiva holística de este tema y que involucre a otros sectores.
- La gobernanza se entiende bajo unos **verbos rectores relacionados con el procesamiento y tratamiento de los datos**. Esta mirada puede limitar la explotación a un número definido de actividades, algo que debe ser repensado en un contexto que constante cambio.

Posteriormente, al iniciar la descripción del plan de acción, esta política señalará una serie de elementos que caracterizarán el marco de gobernanza que se esperaba materializar con la política:

“5.3. Plan de acción”

Las estrategias y acciones de la presente política se regirán bajo los principios que se describen a continuación.

- Los datos son un activo estratégico que genera valor social y económico, por lo que su explotación debe cumplir una función social, orientada a la maximización de beneficios para todos los actores involucrados.
- Los derechos de las personas (naturales y jurídicas) constituyen el límite de la explotación de datos, por lo que el marco jurídico y los instrumentos de gobernanza deben facilitar la generación de valor, garantizando todos los derechos. La dignidad humana orientará la definición de los usos socialmente permitidos de los datos para la generación de valor social y económico.
- Siempre que sea posible, se emplearán mecanismos de colaboración entre el sector público, privado, la academia y la ciudadanía, para maximizar la cooperación entre todos los actores. La articulación entre todos estos guiará las acciones a implementar en la ejecución de la política.”

A partir de esta política se busca entonces generar una gobernanza que permita aprovechar la naciente infraestructura de datos públicos que se está construyendo en el país, respetar los derechos fundamentales y ahondar en la colaboración entre distintos sectores. Como se evidencia de esta política, será necesario entonces establecer mecanismos para definir el valor de los datos, la normativa aplicable a estos y establece la necesidad de generar órganos que faciliten dicha interacción.

Así mismo, otra de las prioridades del documento era el de aumentar la interoperabilidad de los datos en sector público. Frente a este punto se vienen dando importantes avances con la publicación de una versión del Marco de interoperabilidad para Gobierno Digital en agosto de 2019. Esta ha sido definida como:

“la herramienta que acompaña a las entidades en el desarrollo de sus capacidades de intercambio de información, sin importar sus restricciones o su tamaño. Para esto se estructuró dividiendo la interoperabilidad en sus diferentes ámbitos y puntos de vista que constituyen el ejercicio de intercambiar información en cuatro (4) dominios de interoperabilidad.”
(Ministerio de Tecnologías de la Información y las Comunicaciones, 2019).

Muchos de los elementos que caracterizan la gobernanza propuesta en el documento Conpes 3920 están presentes en varios de los casos de estudio a nivel mundial que se describieron ya en este documento y destacados por la OCDE en su informe sobre el tema. Es claro que uno de los principales objetivos de la política era ahondar en los planes de gobierno digital para que respondiera a los criterios desarrollados a nivel internacional y en el cumplimiento de varios de los estándares específicos que la OECD venía promoviendo al respecto.

Por lo tanto, puede señalarse que Colombia ha venido avanzando en los pasos para generar un marco de gobernanza que cumpla con los elementos resaltados a nivel internacional. Sin embargo, todavía quedan dudas sobre el grado de implementación y el impacto que puede llegar a tener esta caracterización en un ecosistema nuevo de políticas públicas que ahondan en la transformación digital y cuarta revolución industrial. A esto se une un nuevo marco institucional que tiene un efecto directo en este tipo de políticas. Así mismo, es importante analizar si este marco tendrá la capacidad de responder a varios de los cambios constantes que experimentan las tecnologías emergentes.





III

UN NUEVO CONTEXTO Y OBJETIVOS DE POLÍTICA PÚBLICA PARA LA GOBERNANZA DE DATOS



A partir del año 2018 se presentan cambios en el ecosistema de tecnología e innovación de Colombia que deben ser considerados en el diseño de un marco de gobernanza de la infraestructura de datos. En primer lugar, surgen nuevos actores dentro de la institucionalidad del país, tales como, la Consejería Presidencial de Innovación y Transformación Digital, hoy conocida como la Consejería Presidencial de Asuntos Económicos y Transformación Digital. Esto implicó un importante cambio en la agenda digital del país, generando un nuevo actor con la posibilidad de coordinar la política pública en el tema y lograr una transversalidad en varios de los temas a tratar, además de asesorar al Presidente de la República directamente en la materia. Esto se ha evidenciado en la promoción de una agenda de transformación digital, con la inclusión de tecnologías emergentes en varios de los proyectos transformacionales que se vienen presentando en el país.

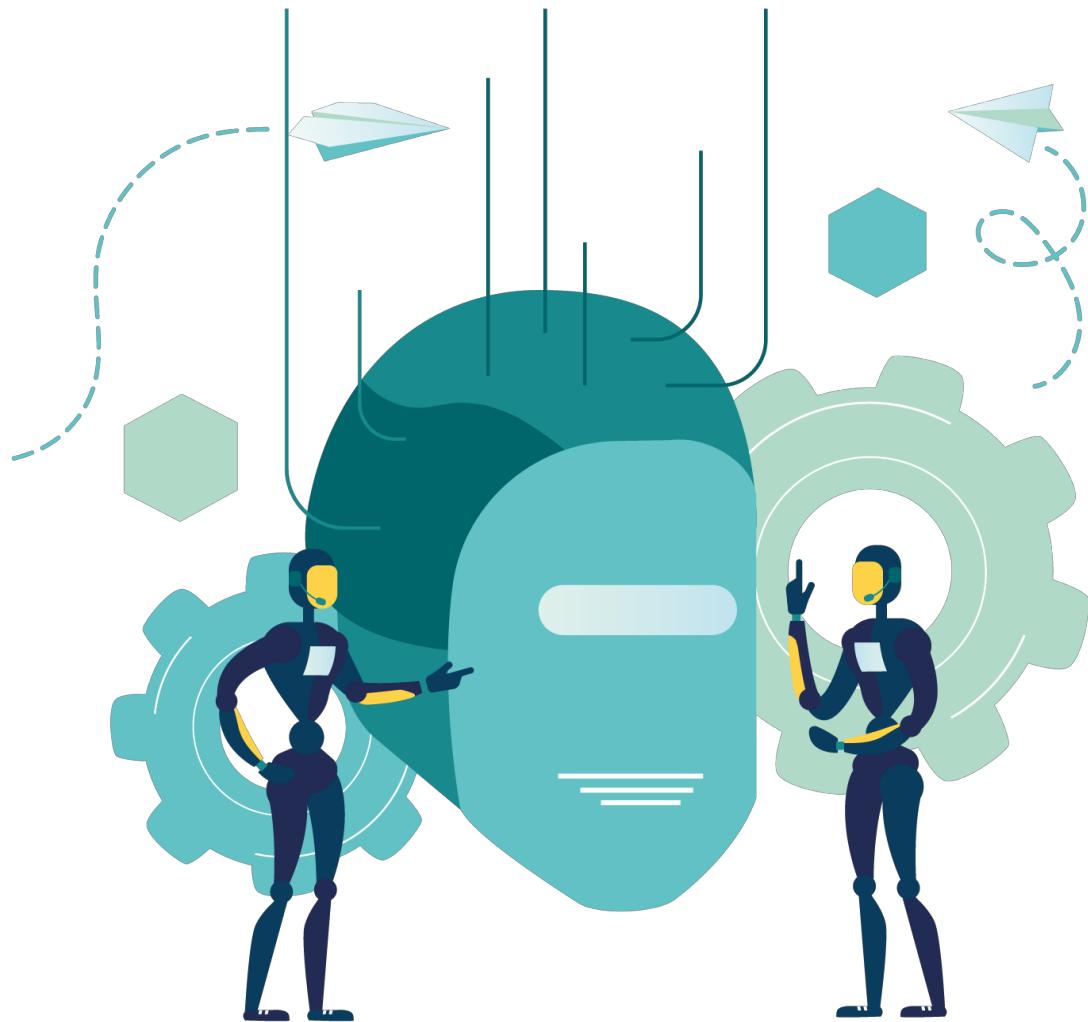
Así mismo, todo esto implica que la política pública en tecnología se centra en el objetivo específico de lograr una transformación digital impulsada por tecnologías como la inteligencia artificial. Esto se manifiesta en la inclusión de artículos sobre transformación digital en el Plan Nacional de Desarrollo (Ley 1955 de 2019), la expedición del documento Conpes 3975 de 2019, además del diseño de un *Framework* de transformación digital del Estado.

De esta forma los objetivos del desarrollo de la infraestructura digital y de datos establecidos en la política pública se dirigen hacia un objetivo específico y es el de transformación digital del Estado de la mano de tecnologías emergentes y propias de la cuarta revolución industrial. Esto incluso ha llevado a proponer cambios institucionales que ahonden en esta postura, como con la propuesta de creación de un Viceministerio de Transformación Digital dentro del Ministerio de Tecnologías de la Información y las Comunicaciones (Ministerio de Tecnologías de la Información y las Comunicaciones, 2020). Uno de los principales objetivos que tendría este Viceministerio es el generar política y planes de implementación de las tecnologías emergentes, siendo una de sus principales tareas generar los lineamientos para la formulación de los planes estratégicos de TI en el proceso de transformación digital de las entidades estatales, el diseño de componentes de tecnología y sus indicadores de seguimiento. Es claro que este sería un actor relevante en la toma de decisiones frente a la configuración de una infraestructura de datos. A esto se une la aparición por primera vez dentro de la institucionalidad del país del Ministerio de Ciencia, Tecnología e Innovación que puede llegar a ser un actor importante en la difusión de datos científicos y de información relevante para el desarrollo tecnológico del país (Fisher & Fortmann, 2010).

Es por esto que no hay duda que el centro de la política pública en tecnología e innovación y de los documentos y normativa que se han construido al respecto están dirigidos a la transformación digital, lo cual implica una priorización a considerar dentro del análisis del marco de políticas e instituciones que ahora se interesan por la infraestructura pública de datos. Esto también significa que se busca priorizar la implementación y despliegue de tecnologías

emergentes como la inteligencia artificial. Entonces el panorama sin duda ha cambiado frente al contexto institucional y de política pública que rodea al documento Conpes de *Big Data*.

Los avances que ha permitido el documento Conpes 3920 de 2018 son valiosos, pero deben ser vistos bajo una nueva perspectiva. Por lo tanto, es necesario establecer la forma como se puede generar un marco de gobernanza que reconozca este nuevo contexto y permita al país ahondar en el cumplimiento de estas políticas, a la vez que genere un marco coherente y que se valga de los esfuerzos ya realizados.





IV HACIA UN NUEVO MODELO DE GOBERNANZA DE DATOS EN COLOMBIA



1 | Una nueva forma de entender el marco institucional y de política pública

Como ya se ha señalado, empieza a ser imprescindible desarrollar una infraestructura de datos que responda a las necesidades propias del diseño, desarrollo e implementación de tecnologías como la inteligencia artificial. De esta forma, el centro de la política pública no debe ser ahora generar una gobernanza que simplemente facilite el acceso e intercambio de datos en un ambiente interoperable. Ahora se busca generar una infraestructura pública que está en constante relación con el marco institucional y regulatorio, con un fin principal: hacer que los datos sean de uso y beneficio común, facilitando una transformación digital que genera un mayor beneficio social. En este caso los temas técnicos deben estar mediados por unas normas y políticas que también brinden las condiciones para el crecimiento de esta infraestructura. Es aquí donde vuelven a surgir conceptos como el de la interoperabilidad legal o normativa que muestre una coordinación que guíe y acompañe el tema técnico.

Para el Centro Berkman Klein de la Universidad de Harvard esto se puede lograr mediante un modelo conocido como *Data Commons Version 1.0*.

El modelo presenta una serie de capas (*layers*) formadas de prácticas organizacionales, instituciones, normas y políticas públicas que se posicionan por encima de las capas propias de la arquitectura técnica y labores de almacenamiento y procesamiento de los datos. Solo reconociendo y privilegiando este orden en las capas es posible para un Gobierno dirigir el diseño técnico de esta infraestructura pública de datos al cumplimiento de objetivos sociales específicos. Es por esto que se presentan tres niveles que determinan cómo las capas de funcionalidad central interactúan con sociedades e instituciones (Berkman Klein Center, 2018). Estos incluyen: las prácticas organizacionales que establezcan los mecanismos específicos para que los *data commons* incentiven la colaboración y la participación multidisciplinaria de distintos actores. En segundo lugar, está un conjunto de instituciones, leyes y políticas que sirven como habilitadores y barreras. Este conjunto determina el grado de accesibilidad de un *data commons*, así como la mitigación de riesgos derivados de la protección de la privacidad y los derechos humanos que puedan verse afectados. Por último, están las personas cuyo conocimiento alimenta el desarrollo y preservación de otras capas y cuya inclusión y educación puede corregir, mejorar y suplementar las capas existentes (Berkman Klein Center, 2018). La adición de estas tres capas establece los *broad data commons*, que permite explorar un amplio rango de aplicaciones y audiencias para la infraestructura pública de datos (Berkman Klein Center, 2018).

Al construirse uno sobre el otro, estos niveles o capas resaltan la importancia de la interoperabilidad (Berkman Klein Center, 2018). Los tipos de interoperabilidad cambian mientras cada uno se mueve más arriba en el marco (Berkman Klein Center, 2018). Dentro de los *narrow data commons*, la interoperabilidad se mueve por estándares técnicos y de datos, mientras que la interoperabilidad semántica y la traducción de conocimiento compartido son dominantes en los *broad data commons* (Berkman Klein Center, 2018). El marco descrito por el Berkman Klein Center también diferencia entre datos de entrada y datos de salida, lo cual permite trazar preguntas y propuestas relacionadas con los datos en las diferentes etapas en el proceso de desarrollo y aplicación de IA (Berkman Klein Center, 2018). Los datos de entrada se refieren a la información utilizada para entrenar un sistema de inteligencia artificial, y los datos de salida se refieren a la información que un sistema de IA pueda generar (Berkman Klein Center, 2018).

El objetivo entonces es facilitar el acceso a los datos a un grupo diverso de actores, generando una infraestructura incluyente y que responda a distintas necesidades. De esta forma, la gobernanza busca no solo hacer un seguimiento técnico a los ciclos de vida de los datos, sino facilitar su uso y potencializar sus beneficios, especialmente en el desarrollo de tecnologías como la inteligencia artificial. Esta misma gobernanza incluso guía la forma como la inteligencia artificial empieza a ser utilizada para labores que generan beneficios sociales, liderando en gran medida la forma como la cuarta revolución industrial va mejorando la calidad de vida de la población. Es aquí donde la gobernanza de datos logra un mayor impacto, que solo determinar la forma como se recolecta, procesa y eliminan los datos:



Lo que hace esta propuesta de *data commons* es involucrar capas adicionales que tienen un efecto social pero que a su vez determinan y guían la infraestructura técnica, los datos y las taxonomías que la definen. Es aquí donde la gobernanza y la institucionalidad dispuesta para este fin se entiende como parte estructural de la infraestructura de datos y del desarrollo de la inteligencia artificial.

2 | Un nuevo modelo técnico (modelo ontológico)

Bajo un mundo de cuarta revolución industrial nos encontramos ante nuevos desafíos en los modelos de gobernanza de datos, especialmente en su componente. Esta obedece ahora a sistemas masivos de información y a datos que interactúan en un contexto específico, propio de sistemas de inteligencia artificial. Esto hace que la gobernanza de datos dentro de las entidades públicas deba cambiar de una postura tradicional y que busca en todo momento una estandarización, hacia el entendimiento de los desafíos que generan las nuevas tecnologías y que caracterizan a las nuevas plataformas que se vienen desarrollando.

Las plataformas de próxima generación, que se construyen como respuesta a datos que se expanden exponencialmente, deben soportar operaciones de negocio, requerimientos regulatorios, automatización de procesos y transformación digital (Ciric, Sohail, & Sharma, 2018). Sin embargo, con el uso de plataformas de próxima generación viene el reto de la escala creciente y la complejidad propia de cada entidad y arquitectura de datos, y subsecuentemente, la gobernanza que lo rodea (Ciric, Sohail, & Sharma, 2018). Al procesar conjuntos grandes de datos personales, datos no estructurados o datos de baja latencia, se vuelve más difícil entender el contexto, importancia y asociaciones de datos y en consecuencia asegurar la confiabilidad y seguridad de los datos (Ciric, Sohail, & Sharma, 2018). Sin un cambio en la mentalidad de gobernanza de datos, las entidades probablemente tendrán dificultades para aprovechar todos los beneficios de negocios de plataformas de próxima generación y gestionar riesgos exitosamente (Ciric, Sohail, & Sharma, 2018).

La gobernanza de datos del futuro no es únicamente sobre maximizar el valor de la efectividad operacional de los datos, la toma de decisiones y los requerimientos regulatorios, sino también de minimizar los riesgos asociados con el manejo pobre de datos (Ciric, Sohail, & Sharma, 2018). Como una capacidad fundamental de gestión de datos, la gobernanza de datos recae sobre cuatro pilares (Ciric, Sohail, & Sharma, 2018):

- **Procesos, políticas, estándares y procedimientos:** incluyen la definición de diferentes niveles de abstracción y de requerimientos de cómo los datos deben gestionarse. Una política de gobernanza de datos debe contener la dirección estratégica de una organización y descripciones de alto nivel, tanto del comportamiento que espera conseguir, como del resultado esperado. Los estándares definen las normas requeridas para asegurar que los datos encajan con el propósito. Los procesos y procedimientos suministran detalles sobre cómo los estándares se ejecutarán. Los estándares típicos incluyen gestión de metadatos, calidad de datos, seguridad de datos, arquitectura de datos y modelaje de datos. Algunos procesos tradicionales de gobernanza de datos no pueden mantener el ritmo con el escalamiento y uso acelerado de plataformas de próxima generación. En consecuencia, estos procesos requieren automatización en el descubrimiento, consumo, interpretación y enriquecimiento de metadatos. En algunos casos, la automatización requiere del uso de tecnologías disruptivas, como tecnologías cognitivas e inteligencia artificial para mantener el ritmo (Ciric, Sohail, & Sharma, 2018).
- **Organización, roles y responsabilidades:** se debe preguntar quién debe responder. Las organizaciones deben definir cuidadosamente órganos formales, roles y responsabilidades dentro de la entidad para facilitar, supervisar y desempeñar procesos dentro del contexto de manejo de datos y gobernanza. Se puede considerar formar un consejo de gobernanza de datos, que maneje el desarrollo e implementación de procesos, políticas, estándares y procedimientos (Ciric, Sohail, & Sharma, 2018).
- **Capacidades de tecnología y herramientas:** las nuevas capacidades de la tecnología incluyen plataformas, herramientas y expertos requeridos para soportar un proceso sólido de gobernanza de datos. Las herramientas típicamente asociadas con la gobernanza de datos son el manejo de metadatos, la calidad de datos, manejo de flujos y herramientas de seguridad de datos, todo lo cual ha cambiado significativamente. Las herramientas actuales tienen una aproximación proactiva y no reactiva a la gobernanza de datos. Adicionalmente, las herramientas tradicionales de gobernanza de datos son limitadas, en tanto que no pueden soportar la variedad de fuentes de metadatos y dispositivos que requieren las plataformas de próxima generación (Ciric, Sohail, & Sharma, 2018).
- **Contenido de metadatos:** se trata de catalogar las características técnicas y operacionales de los datos. El contenido de metadatos es instrumental para gestionar, supervisar y medir los procesos de gobernanza de datos. Las

herramientas de metadatos existentes están más orientadas a los metadatos técnicos y soportan una variedad limitada de activos de metadatos y sus asociaciones mutuas, que deben cambiar. La ontología es un estándar que se puede usar para desarrollar un lenguaje común y hacer referencia a modelos para abordar diferentes componentes de la infraestructura de datos (Ciric, Sohail, & Sharma, 2018).

Es por esto que más que proponer un marco estandarizado de gobernanza de datos, es necesario conocer el nivel de madurez de la infraestructura que viene desarrollando y utilizando cada entidad. Así mismo, es necesario ver cómo cada entidad del sector público viene respondiendo e incorporando sistemas de inteligencia artificial y el lenguaje propio y común que se está generando dentro de los sistemas que implementa cada entidad. Así las cosas, en lugar de disminuir dicha diversidad se debe ahondar en medidas que permitan su aprovechamiento. El marco de gobernanza puede generar lineamientos generales en este fin que le permita a distintas entidades definir su propia gobernanza, reconociendo los objetivos de política pública comunes. Así se empiezan a generar objetos comunes para la gobernanza y mecanismos específicos para que las entidades desarrollen dicha gobernanza, considerando las implicaciones técnicas que se han señalado. En este caso es necesario proteger las ontologías que se vienen generando dentro de la infraestructura de datos, entendiendo la importancia que estas tienen para el despliegue de la inteligencia artificial.

Una ontología es una representación consistente de datos y relaciones de datos dentro de una entidad, un modelo de todos los elementos que se encuentran y conectan varios sistemas de información: los productos y servicios, soluciones y procesos, estructuras organizacionales, protocolos, características del consumidor, métodos de manufactura, conocimiento, contenido y datos de todo tipo. Sin una aproximación consistente y pensada a desarrollar, aplicar y evolucionar una ontología, los sistemas de inteligencia artificial sólo pueden desarrollarse de forma fragmentada, pues les faltaría el fundamento que les permitiría ser suficientemente inteligentes como para tener un impacto. La ontología está en el centro del diseño de información de la inteligencia artificial (Earley & Bernoff, 2020).

Así se logra que las entidades generen herramientas técnicas que les permitan cumplir con objetivos específicos de la política pública como la interoperabilidad de la infraestructura de datos, la transparencia y el acceso de los ciudadanos a ésta.

V
GENERANDO
CONFIANZA EN LA
CREACIÓN DE UNA
INFRAESTRUCTURA
DE DATOS



Otro de los elementos sustanciales que debe caracterizar a una infraestructura de datos y su modelo de gobernanza es la generación de confianza. La generación de confianza debe ser uno de los principales objetivos a considerar y por lo cual se deben diseñar herramientas específicas y concretas que permitan lograr este fin. El paso más efectivo que los gobiernos pueden dar es demostrar a los ciudadanos por qué sus datos importan (Chawda, s.f.). Los gobiernos deben asegurarse de que los ciudadanos entiendan cómo pueden usar sus datos para mejorar sus vidas, que puede ser desde gestionar proactivamente los flujos de datos hasta iluminación de las calles más segura y un mejor uso de energía (Chawda, s.f.). Es necesario que los gobiernos sean transparentes sobre los datos que pueden recolectar y demuestren claramente el valor que resulta de los productos (Chawda, s.f.). La confianza en los datos y la analítica de datos se fundamenta en cuatro claves: (i) calidad: si los datos y modelos analíticos son lo suficientemente buenos; (ii) efectividad: si la analítica entrega los resultados deseados; (iii) integridad: si el uso de datos y analítica es ético y legal, y (iv) resiliencia: qué tan bien se gestionan la gobernanza, seguridad y precisión de los datos en el largo plazo (Chawda, s.f.).

La protección de datos debe estar en todas las etapas, con custodios de datos que tengan un rol clave en asegurar la confiabilidad y tratamiento protegido de los datos (Chawda, s.f.). Estos oficiales van a gestionar los riesgos de la corrupción de datos o las filtraciones de datos (intencionales o no) a lo largo de toda la cadena de valor, que también pueden tener efectos no deseados en la confianza pública (Chawda, s.f.).

Por esto, es necesario ahondar en mecanismos propios para el sector público que generen confianza en la protección de la privacidad, la anonimización y el uso responsable de datos personales. En este punto es necesario que países como Colombia evalúen la inserción de nuevas medidas de protección de datos que sigan aumentando el nivel de confianza en los actores públicos y privados que interactúan con esta información. Por ejemplo, la necesidad de generar mayor claridad en el funcionamiento de las evaluaciones de impacto de tratamiento de datos personales (*privacy impact assessments*) o la privacidad por diseño y por defecto que ya se han venido promoviendo en proyectos específicos como como los de Carpeta Ciudadana Electrónica. Es importante establecer si este tipo de medidas se deben abordar en otras entidades y sectores, con el fin de generar mayor confianza dentro de la ciudadanía en este tipo de proyectos y en la creación de esta infraestructura de datos. Así mismo, se debe explorar la forma como serán evaluadas estas medidas por las autoridades de protección de datos y las exigencias y requisitos que tendrán. Esto debe entrar a complementar el régimen de responsabilidad demostrada que predomina en el país y la necesidad de determinar medidas adecuadas de anonimización, al igual que otras que permitan establecer un análisis de riesgos en el tratamiento de cierta información de forma masiva.

Es importante que un marco de gobernanza permita la evaluación de estos puntos que generan confianza en la ciudadanía frente a la construcción de una infraestructura de datos y en las

modificaciones que debe tener el marco normativo que impacta este punto. En este caso se puede analizar reformas a los decretos que reglamentan la ley 1581 de 2012 y que pueden ahondar en varios de estos puntos.

Es por esto por lo que la determinación de las medidas de protección de datos y privacidad deben ir de la mano y ser coordinadas con los objetivos de política pública que busca generar una infraestructura de datos para el desarrollo de tecnologías como la inteligencia artificial y su despliegue en el sector público y privado. Solo de esta forma se logran objetivos coordinados y el diseño del modelo de gobernanza deben facilitar esta tarea.





VI

NECESIDAD DE
**INTERCAMBIAR MAYOR
INFORMACIÓN CON EL
SECTOR PRIVADO Y
TENER SUS APORTE^S
EN LA CONSTRUCCIÓN
DE LA INFRAESTRUCTURA**



La posibilidad de lograr una interacción mas amplia y constante del sector privado para este fin solo es posible con el diseño de un modelo de gobernanza que así lo permita. Además, la creación de una infraestructura de datos robusta y realmente masiva no puede depender exclusivamente de los datos que administra el sector público. Es claro que los retos que enfrentan hoy en día el Estado y los particulares en un contexto de cambio constante hacen necesaria una mayor colaboración entre estos sectores y estrategias específicas para el intercambio de información.

Por lo tanto, es necesario explorar mecanismos concretos para el intercambio de información entre estos sectores que fortalezca la infraestructura de datos. Distintas medidas de gobernanza de datos pueden servir como instrumentos poderosos para este fin. Por ejemplo, en Holanda, el Standard Business Reporting (SBR) redujo la carga impuesta en los negocios al suministrar información de negocios a las autoridades locales y los bancos (Huyer & Cecconi, 2019). Para este propósito, el SBR definió un marco de gobernanza de datos compartido para datos públicos y privados, que incluía entre otros, un comité dentro del sector público a cargo de definir los objetivos y el programa del SBR y un consejo encargado de decidir el curso de acción, que se beneficia de información suministrada por actores públicos y privados (Huyer & Cecconi, 2019). Adicionalmente, el SBR creó un rol de coordinador para asegurar la implementación coherente del programa, así como una plataforma para que los actores tanto públicos como privados monitoreen y den sus opiniones sobre la implementación del programa (Huyer & Cecconi, 2019). Más aún, el SBR estandarizó las definiciones de datos, utilizando una taxonomía definida por el gobierno holandés y definió procesos de reporte comunes (Huyer & Cecconi, 2019).

El ejemplo de Holanda es solo uno de los múltiples modelos para intercambiar datos entre el sector privado y el gobierno, que no se pueden categorizar como una única aproximación que permita abordar todas las situaciones, sino que los diferentes modelos deben adaptarse dependiendo de cada situación (Huyer & Cecconi, 2019). Se han identificado los siguientes tipos de modelo, que no son los únicos y pueden estar presentes en combinaciones o en híbridos: acuerdo de múltiples partes para compartir datos; donación de datos; colaboración de datos; intermediarios de datos, y compartir datos por regulación (Huyer & Cecconi, 2019).

Otro de estos modelos es el *data trust* (fiducia de datos), que toma el concepto legal de fiducia y lo aplica a los datos (Hardinges, 2018). Históricamente, las fiducias han sido usadas en la ley para conservar activos y tomar decisiones sobre ellos. Un *data trust* toma este concepto de conservar algo y tomar decisiones sobre su uso, pero lo aplica a los datos (Hardinges, 2018). Es una estructura legal que permite la custodia independiente de algunos datos para el beneficio de un grupo de organizaciones o personas, que puede ser la creación de empresas, la investigación de enfermedades o el empoderamiento de comunidades (Hardinges, 2018). En un *data trust*, los fideicomitentes pueden incluir individuos u organizaciones que tienen datos, que otorgan sus derechos a un fiduciario para que controle los datos y tome decisiones sobre los mismos

(Hardinges, 2018). Por su parte, los beneficiarios serían las personas a las que se les otorga acceso a los datos y las personas que se benefician de lo que se cree con esos datos (Hardinges, 2018).

Los fiduciarios tendrían la obligación legal de tomar decisiones sobre los datos únicamente para el interés de los beneficiarios, conforme al deber fiduciario, lo cual podría ayudar a incrementar la confianza de los individuos sobre la forma en que se usan sus datos (Hardinges, 2018). Adicionalmente, los fiduciarios serían responsables por las decisiones que tomen sobre los datos, que siempre deben alinearse con los intereses de los beneficiarios, conforme a la regulación y al marco de gobernanza de datos que les sería aplicable (OECD, 2019). Todo esto también puede ayudar a promover el intercambio de datos en sectores en los que actualmente no se da (OECD, 2019).

En el 2018, el Reino Unido presentó un acuerdo entre el gobierno y la industria de inteligencia artificial, para mantener al país a la vanguardia de la revolución de inteligencia artificial y datos (OECD, 2019). Como parte de este acuerdo, el gobierno se comprometió a explorar mecanismos para el intercambio de datos como *data trusts*, para lo cual se asoció con el Open Data Institute para explorar cómo este mecanismo podría incrementar el acceso a los datos mientras mantiene la confianza (OECD, 2019). Los tres proyectos piloto en los que trabajó esta alianza incluyeron: la lucha contra el tráfico ilegal de vida silvestre, la reducción del desperdicio de alimentos y el mejoramiento de servicios públicos en Greenwich (Office for Artificial Intelligence, 2019). Adicionalmente, se estableció un Consejo de Inteligencia Artificial para cultivar y fomentar una representación más amplia de expertos para enfocarse en temas específicos, que en principio incluyen, pero no se limitan a, ética y datos, adopción, capacidades y diversidad (Office for Artificial Intelligence, 2019).



VII

PROUESTA DE GOBERNANZA PARA LA INFRAESTRUCTURA DE DATOS PÚBLICOS EN COLOMBIA



A partir de los elementos anteriormente analizados, es posible establecer las características de lo que debe ser el marco de gobernanza de la infraestructura de datos, bajo un contexto de una transformación digital jalona por tecnologías como la inteligencia artificial. Así mismo, este marco debe priorizar el diseño de un modelo de decisión en los objetivos de política pública, normativa y regulación que impacta este sistema. Como se verá a continuación la toma de decisiones sobre las características técnicas de esta infraestructura deben responder a una serie de particularidades y deben motivar la creación de un sistema de gobernanza propio de cada entidad. El modelo de gobernanza debe caracterizarse por ser flexible y de fácil capacidad de adaptación y cambio a nuevas tecnologías y retos cambiantes.

Como elemento innovador el modelo gobernanza de datos públicos se caracteriza por tener distintos momentos según alguno de los **cinco objetivos que se quiere desarrollar:**



A continuación, se describen los cinco objetivos que guían el diseño propuesto y la forma como la institucionalidad propuesta responde a cada uno de ellos

1 COORDINACIÓN INSTITUCIONAL

La gobernanza de datos debe generar un marco de coordinación entre distintas entidades, especialmente en el marco de política pública y regulación, más no en las características técnicas de cada sistema. La institucionalidad vigente debe responder a esta necesidad.

Entidad líder: Consejería Presidencial para Asuntos Económicos y Transformación Digital.

Entidades involucradas: Consejería Presidencial para la Competitividad y la Gestión Pública – Privada, Ministerio de Tecnologías de la Información y las Comunicaciones, Ministerio de Ciencia, Tecnología e Innovación, y Departamento Nacional de Planeación.

Consejería Presidencial para Asuntos Económicos y Transformación Digital



Tareas específicas:

En este punto vale la pena anotar que estas entidades deberán reunirse y conjuntamente tomar decisiones sobre el marco regulatorio y las políticas públicas que impactarán la infraestructura de datos. En este caso no conformarán un consejo u órgano de decisión específico, sino que las entidades pertenecientes al organigrama de la Presidencia de la República se encargarán de coordinar el actuar de estas entidades, emitir documentos oficiales y establecer una hoja de ruta específica para el desarrollo de una infraestructura de datos. Los Ministerios representarán el conocimiento sectorizado y las necesidades de los gremios de

tecnología y academia en relación con la creación, procesamiento y acceso a datos. El Departamento Nacional de Planeación brindará conocimiento extenso y recursos técnicos para la formulación de políticas públicas al respecto. El objetivo es que los productos y entregables que se generen dentro de esta coordinación sean publicados como documentos frente a los que distintas entidades líderes e involucradas tienen consenso.

Una de las principales tareas de estas entidades y que debe ser liderada por la Consejería Presidencial para Asuntos Económicos y Transformación Digital es el diseño de un proyecto específico para la infraestructura de datos conocido como “IA y datos públicos”. El objetivo de este proyecto debe ser generar una infraestructura común, conformada por el intercambio de información, bases de datos y aproximaciones a la resolución de problemas para estimular el desarrollo y aplicación de proyectos de inteligencia artificial para el beneficio social (*AI for Good*) (ITU News, 2020). Así mismo, se debe participar en el Global Initiative, que es un marco abierto de colaboración que busca crear un sistema descentralizado para democratizar la solución de problemas con inteligencia artificial (ITU News, 2020).

Esta iniciativa debe estar inspirada en la propuesta impulsada por la Unión Internacional de Telecomunicaciones conocida como ‘AI and Data Commons’ que busca ofrecer recursos para lanzar nuevos proyectos de inteligencia artificial alineados con los Objetivos de Desarrollo Sostenible (ODS) (ITU News, 2020). Esta iniciativa conectará especialistas en inteligencia artificial, usuarios de inteligencia artificial, propietarios de datos y expertos en temas que se beneficiarán de aplicaciones de IA, ayudándoles a alinear los incentivos para innovar y desarrollar soluciones de IA a problemas definidos (ITU News, 2020). El desarrollo y aplicación de inteligencia artificial partirá de infraestructura de última generación, permitiendo que las soluciones de inteligencia artificial escalen con la ayuda de bases de datos compartidas, ambientes de prueba y estimulación, modelos de IA y software asociado y recursos de computación y almacenamiento (ITU News, 2020).

En este caso el liderazgo recae en la Consejería Presidencial para Asuntos Económicos y Transformación Digital, dadas sus características, capacidad técnica y política de coordinación y liderazgo. Así mismo, esta rescata varios de los elementos que buscaba generar el documento Conpes 3920 al momento de proponer un nuevo marco institucional para el desarrollo de la Gobernanza de Datos en Colombia:

“En primer lugar, el reconocimiento de los datos como un activo implica que

los mismos deban gestionarse, planearse y explotarse para generar valor. Para ello, el Departamento Nacional de Planeación, en articulación con el Ministerio de Tecnologías de la Información y las Comunicaciones, publicará en junio de 2019, las condiciones para crear una instancia con capacidad de acción que materialice un nuevo marco de gobernanza para los datos, en articulación con instancias y autoridades existentes, como el Consejo Asesor Nacional de Estadística y otras autoridades con competencias específicas sectoriales, sin afectar las competencias legales de cada una las instancias y autoridades existentes. La instancia propuesta debe superar los esfuerzos anteriores orientados a la creación de comisiones que carecían de herramientas para su operación, así como trascender las actuales capacidades sectoriales que se focalizan, principalmente, en la generación de lineamientos centrados en la gestión de información. En este sentido, dentro de las condiciones de creación de esta instancia, se determinarán los requerimientos para su posterior transformación en un organismo con capacidades robustas para impulsar la transformación digital del sector público, desde una perspectiva transversal, que supere las limitaciones de la intervención sectorial focalizada y dependa del más alto nivel del Gobierno nacional. Las conclusiones del estudio que determine las condiciones de creación de esta instancia serán presentadas a la autoridad competente para lograr su materialización.”

Ya la Consejería cumple con varias de estas características y funciones y tienen una capacidad de coordinación que incluso le ha permitido ser líder de varias de las políticas de mayor importancia en el ecosistema de tecnología e innovación. Esto además evita el desgaste administrativo que implica la creación de una nueva instancia o consejo asesor al respecto. La principal tarea de la Consejería es desarrollar mayores herramientas para fomentar el marco de gobernanza, además de demostrar agilidad y capacidad de adaptación para responder a un ecosistema de tecnología e innovación cambiante. Este tipo de adaptabilidad se dificulta con una institucionalidad poco flexible, rígida y que no responde a esta nueva realidad.



2 | PARTICIPACIÓN DEL SECTOR PRIVADO

En el sistema de Gobernanza de datos debe predominar el intercambio de datos con el sector privado para el bien común y protegiendo la privacidad y derechos de los ciudadanos. Se debe priorizar un modelo gobernanza que facilite la implementación de *data trusts* en el país.

Entidades líderes: Consejería Presidencial para Asuntos Económicos y Transformación Digital y Consejería Presidencial para la Competitividad y la Gestión Pública – Privada y la Autoridad de Protección de Datos Personales de Colombia-Superintendencia de Industria y Comercio.

Entidades involucradas: Ministerio de Tecnologías de la Información y las Comunicaciones, Ministerio de Ciencia, Tecnología e Innovación, Ministerio de Industria Comercio y Turismo y Departamento Nacional de Planeación.



Tareas específicas:

Este es uno de los modelos más innovadores de la gobernanza de datos de Colombia, dado que se busca una coordinación entre la Presidencia de la República y un órgano que ejerce funciones de vigilancia y control para promover una misma política pública de datos. Así mismo, el conocimiento de todas entidades es vital para desarrollar modelos como los de *data trusts* que sean sostenibles y se adapten a las necesidades regulatorias propias de las normas de protección de datos vigentes en Colombia. Así mismo, esto debe responder a las necesidades propias del sector emprendedor y privado, por lo que el liderazgo de Consejería Presidencial para la Competitividad es esencial en este punto.

Así mismo, la Presidencia de la República debe considerar la importancia de este fortalecimiento de la infraestructura de datos con la generación de mecanismos de gobernanza que permitan la participación activa del sector privado, especialmente en el desarrollo de la inteligencia artificial. El modelo del Reino Unido es interesante en este punto, dado que la Oficina de Inteligencia Artificial dentro de la oficina del Primer Ministro viene promoviendo los *data trusts* para facilitar el acceso para el diseño, desarrollo e implementación de esta tecnología, protegiendo los derechos fundamentales.

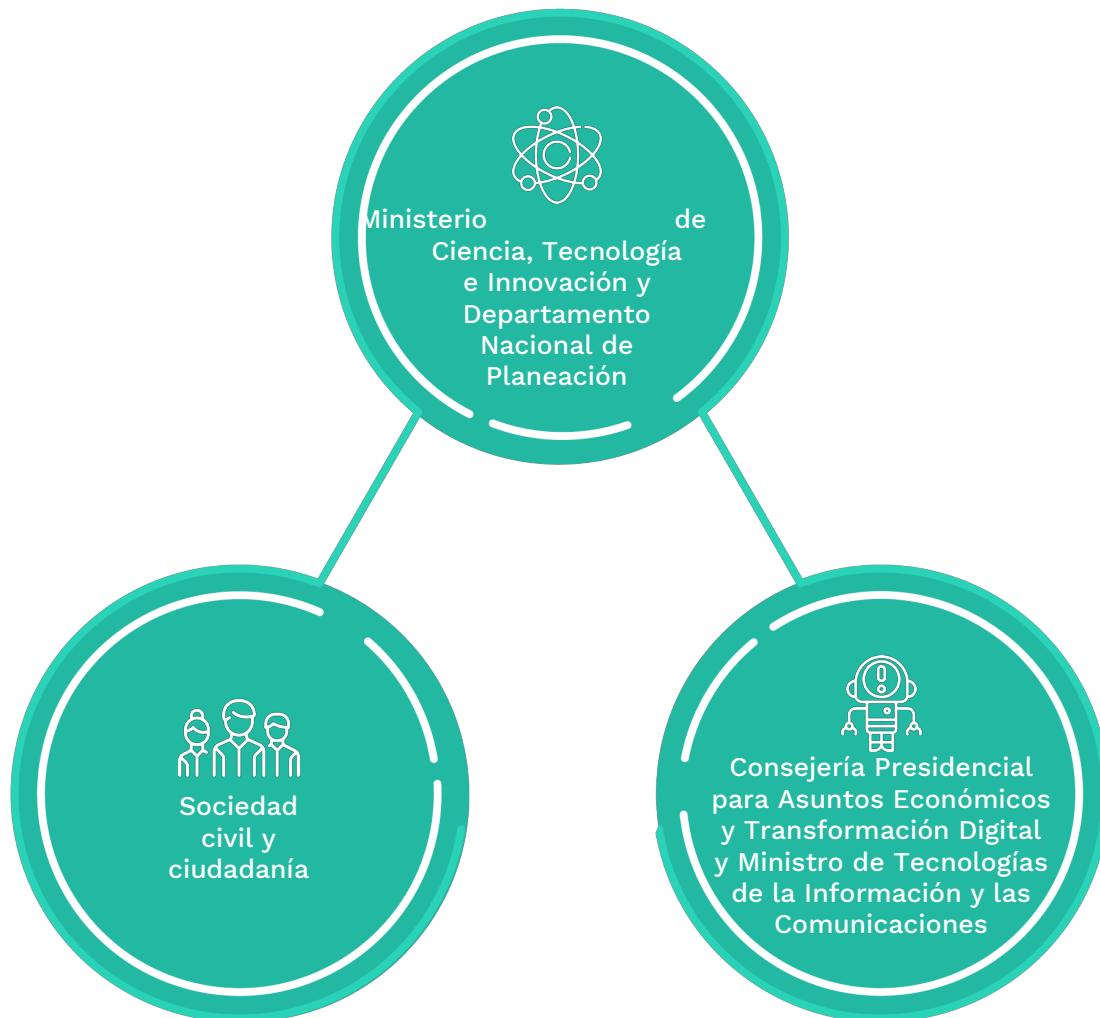


3 | CREACIÓN DE CONFIANZA

Deben existir marcos para la participación de la sociedad civil y la academia en la toma decisiones dentro del marco de gobernanza.

Entidades líderes: Ministerio de Ciencia, Tecnología e Innovación y Departamento Nacional de Planeación.

Entidades involucradas: Consejería Presidencial para Asuntos Económicos y Transformación Digital y Ministerio de Tecnologías de la Información y las Comunicaciones.



Tareas específicas:

Una de las tareas más importantes de este marco de gobernanza es la generación de espacios de diálogo constante con grupos de la sociedad, especialmente los que ejercen actividades de veeduría ciudadana. Así mismo, el sector académico tiene un importante rol al momento de realizar análisis de las políticas y modelos a utilizar e identificar puntos de mejora. Por lo tanto, esta parte del modelo de gobernanza debe ser liderada por el Departamento Nacional de Planeación y el Ministerio de Ciencia, Tecnología e Innovación. La primera entidad tiene amplia experiencia interactuando con la ciudadanía y ha desarrollado metodologías claras para la recepción y análisis de los comentarios y retroalimentación que esta genera, especialmente durante la socialización de los proyectos de documento Conpes.

Por su lado, el Ministerio de Ciencia, Tecnología e Innovación tiene como uno de sus principales propósitos la interacción constante del ecosistema científico y académico con el Gobierno nacional. Por esta razón, se espera que sea una entidad del órgano ejecutivo encargada de dar a conocer las decisiones sobre la infraestructura de datos dentro de esta comunidad e identificar puntos de mejora al respecto.

El objetivo es que estas entidades logren canalizar las opiniones de estos sectores y plasmarlos en las discusiones que tengan con otras entidades como las Consejerías Presidenciales que participan de este modelo de gobernanza u otros ministerios. De esta forma, se logrará una mayor comunicación y atención de este sector de la población ahondando en la generación de confianza frente a la infraestructura que se viene construyendo.



4 | MODELO TÉCNICO (ONTOLÓGICO)

Las entidades que tengan la capacidad de toma de decisión dentro de la gobernanza de datos deben impulsar el uso de un nuevo marco de gobernanza que tenga en cuenta la complejidad de los sistemas de *big data* y otras bases de datos masivas. Estos modelos deben promover y proteger la generación de una infraestructura ontológica y propia de cada entidad.

Entidades líderes: Consejería Presidencial para Asuntos Económicos y Transformación Digital y Ministerio de Tecnologías de la Información y las Comunicaciones.

Entidades involucradas: Archivo General de la Nación, Departamento Administrativo Nacional de Estadística y Departamento Nacional de Planeación.



Tareas específicas:

En este caso se busca que el modelo técnico obedezca a las decisiones de política pública que se han diseñado, discutido e implementando en los primeros modelos. De esta forma, el principal objetivo es que la infraestructura de datos esté respaldada de cuerpo legal y normativo que permita el despliegue técnico adecuado. Por esta razón, este punto lo lidera el Ministerio de Tecnologías de la Información y las Comunicaciones, con el acompañamiento de la Consejería Presidencial para Asuntos Económicos y Transformación Digital. Este Ministerio tiene amplia experiencia desarrollando conocimiento y estándares que sirvan de guía para distintas entidades del estado en la construcción de su arquitectura tecnológica. Cabe resaltar el rol que ha jugado en el desarrollo del modelo de interoperabilidad. Esta ha sido una tarea que ha desarrollado la Dirección de Gobierno Digital y en la que sirve en gran medida su constante relacionamiento con los *Chief Information Officers* de las distintas entidades públicas. En este punto es necesario explorar la forma como el Ministerio puede valerse de herramientas como el CIO para empezar a proponer elementos generales que guíen a estas personas y sus entidades en la creación de sus marcos de gobernanza específicos que sigan los lineamientos de política pública que se han definido de forma coordinada. Esto significa que el ciclo de vida del dato debe facilitar el desarrollo y uso de tecnologías como la inteligencia artificial.

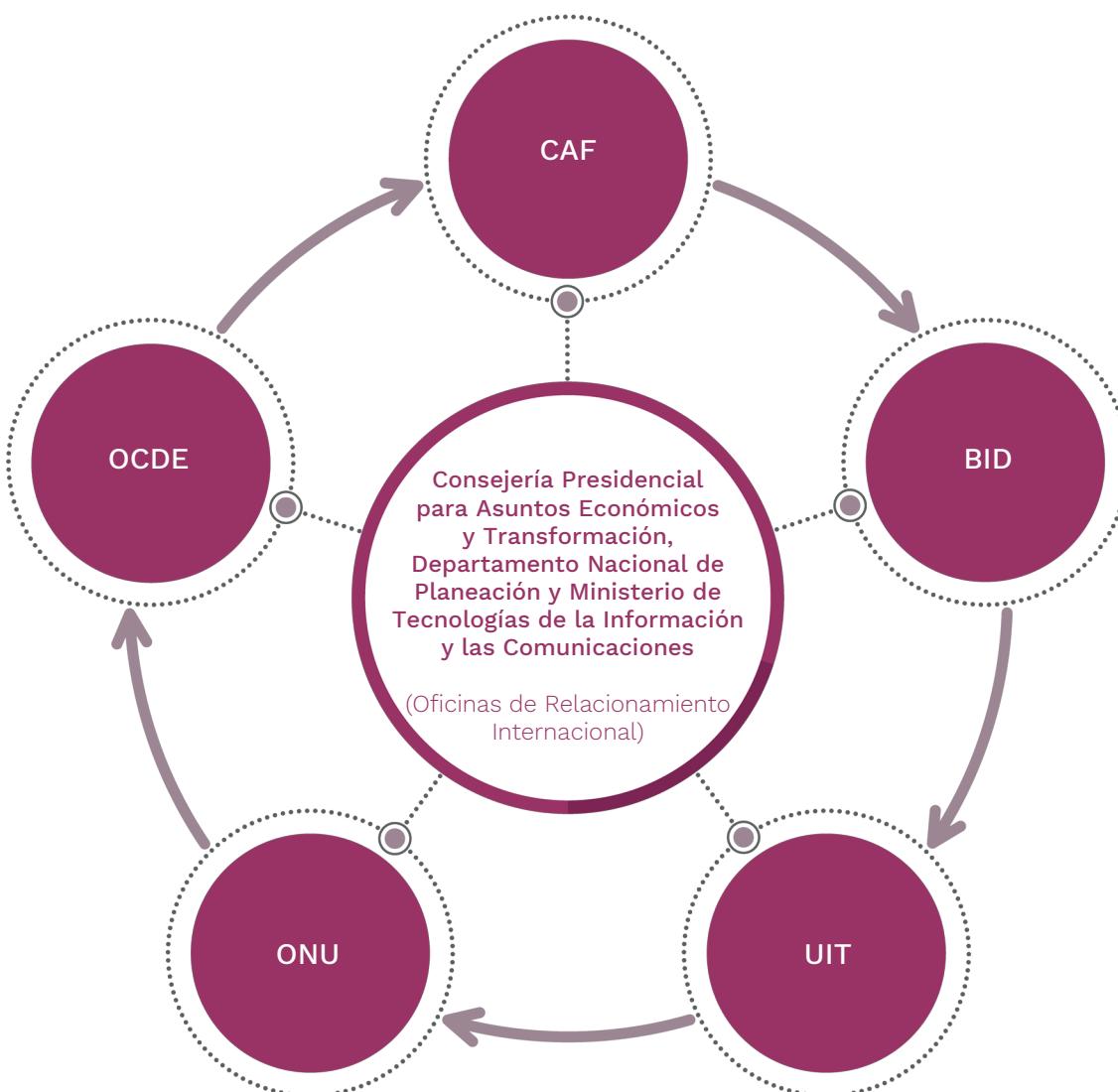
Ahora bien, el proceso de desarrollo de un modelo técnico dentro de cada entidad debe respetar las lógicas propias de cada entidad y la ontología que se viene desarrollando en cada una de estas. Lo importante es que el modelo técnico promueva y fortalezca la identificación y construcción de dichas ontologías, dada la importancia que tienen para el despliegue de nuevas tecnologías emergentes. Es en este punto donde se espera que la gobernanza guíe a las entidades hacia unos mismos principios y objetivos, sin entrar en un detalle técnico que sea poco flexible y adaptable para distintas entidades del Estado.

Por último, la participación en este punto de entidades como el Departamento Administrativo Nacional de Estadística o el Archivo General de la Nación es importante por su experiencia en la creación de estándares y criterios técnicos para la recolección y análisis de la información. Por lo tanto, es esencial que participen de la construcción de un modelo técnico y conozcan y estén comprometidos con los objetivos de política pública para los que se viene gestando un modelo de gobernanza y la infraestructura pública de datos. Es vital que el Ministerio TIC y la Consejería para la Transformación Digital establezcan mecanismos para entrar en contacto con estas entidades y trabajar conjuntamente en el modelo técnico deseado.

5 | IMPACTO INTERNACIONAL

Se debe profundizar en la creación de una infraestructura internacional capaz de interactuar con varios países y entidades en el ecosistema internacional de innovación y tecnología.

Entidades líderes: Consejería Presidencial para Asuntos Económicos y Transformación, Departamento Nacional de Planeación y Ministerio de Tecnologías de la Información y las Comunicaciones.



Tareas específicas:

En este punto se debe profundizar la colaboración con organizaciones multilaterales líderes en la materia para poder generar una infraestructura que sea relevante a nivel internacional. Para este fin se deben generar las siguientes **estrategias**:

- **Recolección de evidencia y casos de estudio:** Colombia debe mostrar proyectos específicos de éxito y uso adecuado de la gobernanza de datos diseñada, promoviéndolos a nivel internacional con el fin de convertir a Colombia en un país atractivo en este sentido. En este caso vale la pena dar difusión a este tipo de casos en iniciativas como la de *AI and Data Commons* propuesto por la UIT. Así mismo, documentar varios de estos casos en inglés es necesario para facilitar dicha difusión y más amplio conocimiento.
- **Adoptar estándares internacionales:** es necesario que Colombia demuestre su adopción de los estándares internacionales y la forma como el modelo de gobernanza se adapta fácilmente a los cambios que se proponen a nivel internacional. Este documento ha buscado proponer un modelo que responda a dichos estándares y a los modelos que se vienen proponiendo e implementando con éxito en otros países.
- **Alcanzar en la infraestructura para el desarrollo de tecnologías emergentes:** este documento también busca proponer un modelo de gobernanza de datos que obedezca directamente a las necesidades de tecnologías emergentes. Solo así se podrá responder a las últimas tendencias internacionales y Colombia y su política de datos no quedará rezagada al respecto.



ORQUESTACIÓN Y MONITOREO

Como se ha podido apreciar el modelo de gobernanza se adapta a distintos momentos y circunstancias específicas. De esta forma, no se busca que exista un único liderazgo y que todo el proceso de deliberación, implementación y diseño de las políticas alrededor de la infraestructura de datos se de dentro de una sola entidad o comité. Ahora bien, es claro que en este caso se requiere de un ejercicio de orquestación que no solo haga posible que estos modelos sean entendidos, sino que funcionen de manera armónica. Por esto es claro que la orquestación de las distintas etapas debe estar en cabeza de la Consejería Presidencial para Asuntos Económicos y Transformación y el Ministerio de Tecnologías de la Información y las Comunicaciones. Estas entidades tienen el rol político y técnico necesario para desarrollar la tarea de liderar este marco de gobernanza y tomar decisiones frente a la forma como se generará una infraestructura que responda a las necesidades de nuevas tecnologías emergentes. Así mismo, su capacidad de coordinación permitirá no solo lograr una interoperabilidad de los datos, sino de las normas, regulaciones e iniciativas que se dan dentro del Estado Colombiano alrededor de la infraestructura de datos. Así mismo, esto evita la generación de un consejo o comité que puede llegar a ser un ejercicio más desgastante en términos de tiempo y recursos, y que no se adapte fácilmente a los múltiples cambios que se pueden ir presentando.

Por último, vale la pena anotar que un modelo de gobernanza exitoso requiere de mecanismo y herramientas de monitoreo constante. Es en este punto donde el Departamento Nacional de Planeación puede prestar un enorme servicio dadas las herramientas que ya tiene de seguimiento y los canales de interacción que ya ha desarrollado con múltiples entidades del Estado. Vale la pena anotar que además esta ya no es solo una tarea propia del documento Conpes 3920 de 2018, sino también del documento Conpes 3975 de 2019 por lo que valdría la pena desarrollar dentro del seguimiento de esta política indicadores específicos para este fin. El Departamento Nacional de Planeación puede proveer información constante a la Consejería Presidencial para Asuntos Económicos y Transformación Digital y el Ministerio de Tecnologías de la Información y las Comunicaciones frente a la forma como se vienen impulsando las medidas, objetivos y políticas frente a la infraestructura de datos que viene promoviendo. Así mismo, el Ministerio TIC y la Consejería puede ahondar en este punto dentro del seguimiento que realiza a las entidades del Estado dentro de los proyectos transformacionales y del trabajo que realizan los CIO. Igualmente, la Consejería Presidencial para Asuntos Económicos y Transformación Digital puede aprovechar varias de las herramientas de seguimiento que se vienen desarrollando dentro de la Consejería Presidencial para la Gestión y el Cumplimiento, con el fin de hacer un seguimiento específico de la implementación de medidas que impactan directamente la construcción de la infraestructura de datos y que permitan medir la eficiencia del sistema de gobernanza que están implementando.

**Referencias:**

- Berkman Klein Center. (2018, Junio 21). Data Commons Version 1.0: A Framework to Build Toward AI for Good. Retrieved from Medium: <https://medium.com/berkman-klein-center/data-commons-version-1-0-a-framework-to-build-toward-ai-for-good-73414d7e72be>
- Chawda, V. (n.d.). Building trust in government's use of data. Retrieved from KPMG: <https://home.kpmg/xx/en/home/insights/2018/06/building-trust-in-governments-use-of-data.html>
- Ciric, B., Sohail, O., & Sharma, P. (2018). Data governance for next-generation platforms. Retrieved from Deloitte: <https://www2.deloitte.com/content/dam/Deloitte/us/Documents/technology/us-big-data-governance.pdf>
- Earley, S., & Bernoff, J. (2020, Abril 28). Is Your Data Infrastructure Ready for AI? . Retrieved from Harvard Business Review: <https://hbr.org/2020/04/is-your-data-infrastructure-ready-for-ai>
- e-estonia. (2018). interoperability services x-road. Retrieved from e-estonia: <https://e-estonia.com/solutions/interoperability-services/x-road/>
- Fisher, J., & Fortmann, L. (2010). Governing the data commons: Policy, practice, and the advancement of science. *Information & Management*, 47(4), 237-245.
- Gasser, U. (2015). Interoperability in the Digital Ecosystem. *SSRN Electronic Journal*.
- Hardinges, J. (2018, Octubre 19). Defining a 'data trust'. Retrieved from Open Data Institute: <https://theodi.org/article/defining-a-data-trust/>
- Huyer, E., & Cecconi, G. (2019, Marzo). Analytical Report 12: Business-to-Government Data Sharing. Retrieved from European Data Portal: europeandataportal.eu/sites/default/files/analytical_report_12_business_government_data_sharing.pdf
- ITU News. (2020, Abril 23). Join the Global Initiative on 'AI and Data Commons'. Retrieved from ITU News: <https://news.itu.int/join-global-initiative-ai-data-commons/>
- Ministerio de Tecnologías de la Información y las Comunicaciones. (2019, agosto). Marco de interoperabilidad para Gobierno Digital. Retrieved from Ministerio de Tecnologías de la Información y las Comunicaciones: http://lenguaje.mintic.gov.co/sites/default/files/archivos/marco_de_interoperabilidad_para_gobierno_digital.pdf
- Ministerio de Tecnologías de la Información y las Comunicaciones. (2020). Proyecto de Decreto "Por el cual se modifica la estructura del Ministerio de Tecnologías de la Información y las Comunicaciones, se determinan las funciones de sus dependencias y se dictan otras disposiciones". Retrieved from Ministerio de Tecnologías de la Información y las Comunicaciones: https://mintic.gov.co/portal/604/articles-126484_pd_modifica_estructura_mintic.pdf
- OECD. (2017, Septiembre 1). Digital Government Review of Norway. Retrieved from OECD: https://www.oecd-ilibrary.org/governance/digital-government-review-of-norway_9789264279742-en



OECD. (2019). 2. Data governance in the public sector . In OECD, The Path to Becoming a Data-Driven Public Sector. París: OECD Publishing.

Office for Artificial Intelligence. (2019, Abril). AI Sector Deal One Year On. Retrieved from Office for Artificial Intelligence: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/819331/AI_Sector_Deal_One_Year_On__Web_.pdf

Republic of Estonia Information System Authority. (2019, Julio 9). Data Exchange Layer X-tee. Retrieved from Republic of Estonia Information System Authority: <https://www.ria.ee/en/state-information-system/x-tee.html>

Sweeney, K. (2020). An operational Data Governance Framework for New Zealand Government. Retrieved from Stats NZ: <https://statsnz.contentdm.oclc.org/digital/collection/p20045coll1/id/2657>

The British Academy and the Royal Society. (2017, Junio). Data management and use: Governance in the 21st century. Retrieved from The Royal Society: <https://royalsociety.org/~media/policy/projects/data-governance/data-management-governance.pdf>



MODELO DE GOBERNANZA DE LA **INFRAESTRUCTURA DE DATOS** PARA EL DESARROLLO DE TECNOLOGÍAS EMERGENTES



El futuro
es de todos

Gobierno
de Colombia