



El futuro digital  
es de todos

MinTIC

# Convocatoria Big Data

Términos de referencia

PREPARADO POR: Carlos Julio León Caicedo  
Febrero 2021

# Data Sandbox

## **Ministerio de Tecnologías de la Información y las Comunicaciones**

Karen Abudinen Abuchaibe

### **Ministra de Tecnologías de la Información y las Comunicaciones**

Germán Camilo Rueda Jiménez

### **Viceministro de Transformación Digital**

Aura María Cifuentes

### **Directora de Gobierno Digital**

Susan Simoneth Suarez Gutiérrez

### **Subdirectora de Fortalecimiento de Capacidades Públicas Digitales**

Luisa Fernanda Medina

Thiago Grijo Dal Toe

Carlos Julio León Caicedo

### **Equipo de trabajo – Gobierno Digital**

## **Desarrollo técnico del documento**

El presente documento se ha desarrollado en el marco del proyecto Data Sandbox (Espacio colaborativo para el desarrollo de proyectos piloto de analítica y Big Data por parte de entidades del sector público).

Jairo David Acosta Fernández – Gerente de proyectos

Luis Alfonso Patarroyo – Experto en Big Data

### **Unión Temporal Nube Pública 2019**

<b>Versión</b>	<b>Observaciones</b>
Versión 1 Febrero 2021	<b>Título</b> Términos de referencia para la convocatoria a la postulación de proyectos piloto de Analítica y Big Data por parte de entidades públicas del orden nacional o territorial a desarrollarse en el Data Sandbox.

Comentarios, sugerencias o correcciones pueden ser enviadas al correo electrónico:

[gobiernodigital@mintic.gov.co](mailto:gobiernodigital@mintic.gov.co)

Términos de referencia de la convocatoria.



Este documento de la Dirección de Gobierno Digital se encuentra bajo una [Licencia Creative Commons Atribución 4.0 Internacional](https://creativecommons.org/licenses/by/4.0/).

# Contenido

1. Condiciones generales.....	4
2. Acerca de la convocatoria y la inscripción .....	5
3. Acerca de las temáticas de participación .....	6
4. Acerca de los criterios de evaluación .....	9
4.1. Criterio 1: Dimensiones Básicas del Proyecto .....	9
4.2. Criterio 2: Conformación de equipo de trabajo de la Entidad .....	9
4.3. Criterio 3: Duración del Proyecto Piloto.....	11
4.4. Criterio 45: Habilidades del equipo en herramientas de Big Data .....	11
5. Acerca de la generación de capacidades.....	12
6. Acerca de la entrega de los servicios de Nube .....	13
7. Contrapartida de las Entidades .....	13
8. Cronograma de la convocatoria .....	14
9. Propiedad intelectual .....	15
11. Anexos .....	16

# 1. Condiciones generales

**Apreciado Usuario:** la convocatoria “Convocatoria Big Data al servicio de las entidades públicas” (en adelante la Convocatoria) disponible en el siguiente enlace: [Convocatoria Big Data](#), surge como respuesta a la necesidad de acceso a infraestructura, mejora de habilidades, transferencia de conocimiento que las entidades públicas nos han planteado para mejorar sus capacidades en el uso y aprovechamiento de datos estructurados, semiestructurados y no estructurados para el desarrollo de soluciones a problemáticas públicas y/o ciudadanas a través del uso de tecnologías de Big Data. Algunas de las entidades públicas del orden nacional (DANE, DNP, UPRA, Contraloría, FNA, entre otras), han avanzado en la conformación de unidades, equipos, direcciones y/o áreas orientadas al desarrollo de proyectos de Analítica y Big Data y el uso de las tecnologías de ciencias de datos. Adicionalmente, las entidades públicas del país, son uno de los actores que tienen una gran importancia para las iniciativas de [datos abiertos](#) y [Data Sandbox](#) (Espacio colaborativo para el desarrollo de proyectos de analítica y Big Data por parte de entidades del sector público), teniendo en cuenta que estas entidades por su naturaleza, consolidan información a nivel país, cuentan con equipos de trabajo con niveles de conocimiento intermedios y/o avanzados en el uso de tecnologías de Big Data y su quehacer diario son un escenario natural para el mayor aprovechamiento y la transformación de sus datos en nuevas soluciones para el sector público.

Los presentes términos y condiciones vinculan a cualquier persona natural o jurídica que ingrese en este sitio web y registre su propuesta. El Ministerio de Tecnologías de la Información y las Comunicaciones (En adelante el MINTIC) solicita al Usuario de esta página, que lea detallada y juiciosamente estos términos y condiciones de participación (en adelante los Términos de Referencia), antes de iniciar su exploración o utilización o registro. El MINTIC asumirá que los Términos son aceptados, cuando el participante realice su registro correspondiente antes del inicio de la convocatoria.

La utilización de este sitio implica la aceptación expresa y plena de las condiciones generales de acceso y uso de la versión publicada, en el instante en que el usuario acceda al mismo, sin perjuicio de las circunstancias particulares que pudieran aplicarse a algunos contenidos o servicios en concreto.

Si el Usuario no está de acuerdo con estos Términos de Referencia o con cualquier disposición de la Política de Privacidad, le sugerimos que se abstenga de acceder o navegar por este Sitio.

El MINTIC a través de la Dirección de Gobierno Digital, podrá incorporar cambios en los Términos y Condiciones según considere oportuno y no se hará responsable por cualquier daño, pérdida, gasto, perjuicio o decepción causada por haber participado en la convocatoria.

El MINTIC se reserva la facultad de decisión respecto a toda situación no prevista en las bases de la convocatoria. Las decisiones que se tomen serán definitivas en el transcurso de la convocatoria.

El MINTIC no será responsable de errores técnicos en las redes de telecomunicaciones, acceso a internet o de otro tipo, que impidan la participación en la convocatoria.

El MINTIC no se hará responsable por cualquier costo asociado o que no estén específicamente incluidos en los incentivos de asignación de infraestructura en el Data Sandbox, Capacitaciones, Talleres y/o Acompañamiento de Expertos en Big Data a otorgar a las entidades beneficiarias.

## 2. Acerca de la convocatoria y la inscripción

Esta convocatoria está dirigida a las **entidades públicas del orden nacional y territorial de Colombia**, cuyo objetivo es el de recibir la postulación de **proyectos piloto de analítica y Big Data** a desarrollarse en el **Data Sandbox**, el cual, es un espacio colaborativo para que las entidades públicas del país puedan realizar diferentes proyectos piloto de Analítica y Big Data. La plataforma puede ser empleada en la exploración de conjuntos de datos de manera colaborativa e interactiva entre varios Usuarios del Data Sandbox, con el fin de probar, experimentar, ensayar y determinar soluciones aplicables en el contexto real a problemáticas públicas y/o ciudadanas. Este espacio colaborativo cuenta con altas capacidades para el almacenamiento y procesamiento de datos estructurados, no estructurados y semiestructurados, a través del uso de tecnologías de Big Data. A continuación se deben tener en cuenta las siguientes recomendaciones para la inscripción y postulación de equipos de trabajo y proyectos:

- Las inscripciones se deben hacer únicamente a través del formulario dispuesto para tal fin en la página <http://bit.ly/3eLMI8y>.
- Se pueden inscribir equipos conformados por funcionarios y/o contratistas **de una entidad pública del orden nacional o territorial de Colombia**. Cada equipo de ciencia de datos deberá conformarse con los siguientes perfiles interdisciplinarios, y/o las personas que conformen estos equipos deberán asumir uno o más de los siguientes roles o perfiles: Responsable líder, Científico de datos, Ingeniero de Datos, Analista de datos, Artista de datos, Estadístico y/o Administrador de bases de datos. Muchos de los roles y perfiles descritos anteriormente, pueden ser asumidos por una sola persona, de igual manera pueden existir nuevos roles o perfiles que las mismas Entidades definan y que complementen al equipo de trabajo. Para ampliar esta información por favor leer la sección **“Conformación de equipo de trabajo de la Entidad”** de los [Términos y condiciones de uso](#) del Espacio colaborativo para el desarrollo de proyectos piloto de Analítica y Big Data (Data Sandbox).
- Los proyectos piloto tendrán una **duración de entre uno y cinco meses** para su desarrollo en el Data Sandbox.
- Los equipos de trabajo de las entidades deberán conformarse cómo mínimo con dos personas y hasta siete personas. Se sugiere cómo máximo siete personas, pero las entidades son autónomas en la conformación de los equipos de trabajo, de igual manera se les dará acceso al espacio asignado a la entidad en el Data Sandbox a los perfiles que interactúan con los servicios de Big Data.
- A través de la convocatoria se permitirá el ingreso al Data Sandbox de hasta **ocho (8) proyectos**, estos proyectos y los equipos ganadores accederán a **un espacio en el Data Sandbox**, adicionalmente recibirán el **acompañamiento técnico** por parte de un experto en Big Data que solucionará inquietudes técnicas causadas por el uso de los ambientes dispuestos para que las entidades piloteen sus proyectos, de igual manera accederán a **capacitaciones y talleres de formación** en arquitecturas y servicios de nube, que para el caso del Data Sandbox operan bajo el segmento de nube pública de Microsoft.
- Para la postulación de los proyectos se deberá diligenciar la siguiente [plantilla de proyecto](#), la cual será presentada por parte del equipo al comité evaluador de la convocatoria de acuerdo con la sección del **cronograma de la convocatoria** descrita en este documento.

### 3. Acerca de las temáticas de participación

Los proyectos que hagan uso del espacio colaborativo Data Sandbox en el marco de la convocatoria, podrán alinearse con algunas de las siguientes temáticas: Objetivos de Desarrollo Sostenible, metas del Plan Nacional de Desarrollo vigente y/o los objetivos estratégicos de las entidades públicas del orden nacional.

## **OBJETIVOS DE DESARROLLO SOSTENIBLE**



Fuente: <https://www.un.org/sustainabledevelopment/es/2015/09/la-asamblea-general-adopta-la-agenda-2030-para-el-desarrollo-sostenible/>

# 20

**METAS**

del Plan Nacional de Desarrollo  
que transformarán a COLOMBIA



1

Llegar a **2 millones de niños** con educación inicial: aumento del **67%**.



Pasar de **5,3 millones** de niños a **7 millones** en el PAE.

2

Duplicar los **estudiantes** en jornada única en colegios oficiales: de **900 mil** a **1,8 millones**.



3

Fortalecimiento a **las 61 IES públicas**. Avance gradual en gratuidad para **320.000 jóvenes**, reconocimiento a la excelencia. **Aumento de cobertura** de **53% al 60%**.



4

Multiplicar más de **cuatro veces** Jóvenes en Acción, llegando a **500 mil cupos**.



5

**600 mil** beneficiarios de mejoramiento de vivienda y **520 mil VIS** iniciadas.



6

**Saneamiento de deuda por recobros** del régimen contributivo de salud a **31 de diciembre de 2019**.



7

**550 mil** productores con asistencia técnica agro y **300 mil** con agricultura por contrato.



8

Lograr el **60%** de actualización catastral, frente al **20%** de hoy. Casi **duplicar la velocidad** de titulación.



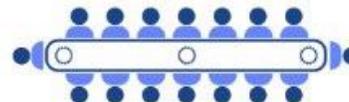
9

Apoyar a **4.000** empresas con fábricas de productividad, frente a **200** actuales.



10

Crear **1,6 millones de empleos** y reducir el desempleo de **9,4%** a **7,9%**, el más bajo desde los **90s**.



Fuente: [https://twitter.com/dnp\\_colombia/status/1094688978090016768?lang=pt](https://twitter.com/dnp_colombia/status/1094688978090016768?lang=pt)

# 20



## **METAS** del Plan Nacional de Desarrollo que transformarán a COLOMBIA

**11** Erradicar **280 mil** hectáreas de cultivos ilícitos.

**12** Detener el crecimiento de la **deforestación** en 2022: hoy crece al **23% anual**.

**13** Llevar la tasa de homicidios a su **menor nivel** en **27 años**: **23,2** por cada 100 mil habitantes, frente a **25,8** hoy.

**14** Lograr que **11,8 millones** de hogares (**70%**) estén conectados a internet: hoy lo están **7,4 millones (50%)**.

**34** trámites de alto impacto ciudadano, transformados digitalmente.

**15** Mayor dinámica de los sectores de economía naranja: crecimiento real de **2,9%** a **5,1%** en cuatro años.

**16** **Duplicar** la inversión pública y privada en ciencia y tecnología a **1,5% del PIB**.

**17** Aumentar en más del **doblo** la **red férrea** en operación comercial: llegar a **1.077 Km**.

**18** Aumentar **capacidad** de generación con energías limpias en **1.500 MW**, frente a **22,4 MW** hoy.

**19** Sacar a **1,5 millones** de personas de la **pobreza extrema**.

**20** Sacar a **2,9 millones** de personas de la **pobreza**.

Fuente: [https://twitter.com/dnp\\_colombia/status/1094688978090016768?lang=pt](https://twitter.com/dnp_colombia/status/1094688978090016768?lang=pt)

## 4. Acerca de los criterios de evaluación

Para la selección de los proyectos a implementarse en el espacio colaborativo (Data Sandbox), se realizará una selección de los **ocho (8) mejores proyectos**, con los **puntajes más altos iguales o superiores al 70% en los resultados de evaluación de cada proyecto**, para ello se tendrán en cuenta los siguientes criterios de evaluación para esta priorización:

### 4.1. Criterio 1: Dimensiones Básicas del Proyecto

Este criterio tiene un peso total del 30% para la evaluación del proyecto, el criterio se divide en los subcriterios de Volumen (60%) y Variedad (40%), a continuación la descripción de los subcriterios:

Subcriterio	Descripción	Tipos de Evidencias	Puntajes asignados
Volumen	Describe el volumen de los datos que serán procesados en el Data Sandbox, es importante especificar la cantidad máxima de datos que el proyecto soportará.	Especificar los siguiente: <ul style="list-style-type: none"> <li>• Tamaño aproximado de la información.</li> <li>• Tipo(s) de fuente(s) Ej. transacciones online, datos científicos y de investigación, datos de sensores, redes sociales.</li> <li>• Nombre de la(s) fuente(s).</li> </ul>	<ul style="list-style-type: none"> <li>• Mayor a 1TB y hasta 2TB: 1</li> <li>• Mayor a 500GB y hasta a 1TB: 0.7</li> <li>• Mayor a 300GB y hasta 500GB: 0.4</li> <li>• Menor o igual a 300GB: 0.2</li> </ul>
Variedad	La variedad de los datos se refiere a los múltiples formatos y tipos de datos que deben ser compatibles con las soluciones de Big Data.	Especificar los tipos de datos que apliquen: <ul style="list-style-type: none"> <li>• Datos estructurados (BD relacionales, data tabular).</li> <li>• Datos no estructurados (Imágenes, Video, Audio).</li> <li>• Datos semiestructurados (XML, JSON, Sensores).</li> </ul>	<ul style="list-style-type: none"> <li>• Datos no estructurados (Imágenes, Video, Audio): 1</li> <li>• Datos semiestructurados (XML, JSON, Sensores): 0.8</li> <li>• Datos estructurados (BD, data tabular): 0.5</li> </ul>

### 4.2. Criterio 2: Conformación de equipo de trabajo de la Entidad

Este criterio tiene un peso total del 40% para la evaluación del proyecto, el criterio se divide en los subcriterios de Responsable líder (5%), Científico de datos (25%), Ingeniero de Datos (25%), Analista de datos (10%), Artista de datos (5%), Estadístico (20%) y Administrador de bases de datos (10%), a continuación la descripción de los subcriterios:

Subcriterio	Descripción	Tipos de Evidencias	Puntajes asignados
Responsable líder, Científico de datos, Ingeniero de Datos, Analista de datos, Artista de datos, Estadístico, Administrador	Las Entidades que deciden enfrentarse a estos nuevos retos de aplicar la disciplina de Big Data en el desarrollo de sus proyectos piloto deberán contemplar la conformación de un equipo de trabajo o equipo de ciencia de datos, con un mínimo de	Adjuntar la siguiente documentación a través de certificaciones: <ul style="list-style-type: none"> <li>• Responsable líder – Certificar que es funcionario de planta o contratista de la entidad postulante.</li> <li>• Científico de datos – Certificar experiencia (Nombrar proyectos donde ha participado) en el manejo de modelos predictivos,</li> </ul>	Responsable líder: <ul style="list-style-type: none"> <li>• Funcionario de Planta: 1</li> <li>• Contratista: 0.5</li> </ul> Científico de datos: Manejo de modelos predictivos y experiencia en lenguajes como R, Python, SQL, Scala, entre otros. <ul style="list-style-type: none"> <li>• Experiencia mayor a 2 años: 1</li> <li>• Experiencia entre 1 a 2 años: 0.7</li> <li>• Experiencia menor a 1 año: 0.3</li> </ul>

Subcriterio	Descripción	Tipos de Evidencias	Puntajes asignados
de bases de datos	dos y hasta siete integrantes, no obstante, la Entidad podrá disponer de más miembros en el equipo.	<p>en trabajos donde ha utilizado lenguajes como R, Python, SQL o Scala. Si ha realizado cursos relacionados con la ciencia de datos añadir los certificados respectivos.</p> <ul style="list-style-type: none"> <li>Ingeniero de Datos - Certificar experiencia (Nombrar proyectos donde ha participado) en el manejo de servicios y herramientas para la ingesta, la salida y la transformación de datos procedentes de múltiples orígenes. Si ha realizado cursos relacionados con la ingeniería de datos añadir los certificados respectivos.</li> <li>Analista de datos - Certificar experiencia (Nombrar proyectos donde ha participado) en proyectos donde ha desempeñado este rol.</li> <li>Artista de datos - Certificar experiencia (Nombrar proyectos donde ha participado) en proyectos donde ha desempeñado este rol.</li> <li>Estadístico - Certificar experiencia (Nombrar proyectos donde ha participado) en el manejo de métodos estadísticos aplicados a grandes volúmenes de datos (análisis descriptivo - análisis predictivo) y en habilidades en lenguajes como R, Python, SQL o Scala. Si ha realizado cursos relacionados con el análisis de datos añadir los certificados respectivos.</li> <li>Administrador de bases de datos - Certificar experiencia (Nombrar proyectos donde ha participado) en el manejo de herramientas y tecnologías que se usan para almacenar y gestionar grandes cantidades de datos.</li> </ul>	<p>Ingeniero de Datos: Habilidades en el uso de servicios y herramientas para la ingesta, la salida y la transformación de datos procedentes de múltiples orígenes.</p> <ul style="list-style-type: none"> <li>Experiencia mayor a 2 años: 1</li> <li>Experiencia entre 1 a 2 años: 0.7</li> <li>Experiencia menor a 1 año: 0.3</li> </ul> <p>Analista de datos: Experiencia en recopilación y análisis de grandes volúmenes de datos de forma estadística.</p> <ul style="list-style-type: none"> <li>Experiencia mayor a 2 años: 1</li> <li>Experiencia entre 1 a 2 años: 0.7</li> <li>Experiencia menor a 1 año: 0.3</li> </ul> <p>Artista de datos: Experiencia en representar de manera sencilla (gráficos, infografías y herramientas visuales), los resultados del análisis de los datos para comprender grandes volúmenes de información.</p> <ul style="list-style-type: none"> <li>Experiencia mayor a 2 años: 1</li> <li>Experiencia entre 1 a 2 años: 0.7</li> <li>Experiencia menor a 1 año: 0.3</li> </ul> <p>Estadístico: Entre otras tener experiencia en lenguajes como R, Python, SQL o Scala.</p> <ul style="list-style-type: none"> <li>Experiencia mayor a 2 años: 1</li> <li>Experiencia entre 1 a 2 años: 0.7</li> <li>Experiencia menor a 1 año: 0.3</li> </ul> <p>Administrador de bases de datos: Estar familiarizado con las nuevas herramientas y tecnologías que se usan para almacenar y gestionar grandes cantidades de datos.</p> <ul style="list-style-type: none"> <li>Experiencia mayor a 2 años: 1</li> <li>Experiencia entre 1 a 2 años: 0.7</li> <li>Experiencia menor a 1 año: 0.3</li> </ul>

#### 4.3. Criterio 3: Duración del Proyecto Piloto

Este criterio tiene un peso total del 10% para la evaluación del proyecto, el criterio se divide en el subcriterio de Duración del proyecto, a continuación la descripción del subcriterio:

Subcriterio	Descripción	Tipos de Evidencias	Puntajes asignados
Duración del proyecto	Los pilotos deberían tener una duración máxima de 4 meses, aunque por experiencia de otros pilotos realizados pueden oscilar entre 3 y 5 meses.	Especificar el tiempo que se utilizará en implementar y publicar los resultados del proyecto piloto propuesto. Diligenciar la plantilla de postulación del proyecto - <a href="https://sandbox.datos.gov.co/app/img/ppt/PlantillaPostulacionProyectosPiloto.ppt">https://sandbox.datos.gov.co/app/img/ppt/PlantillaPostulacionProyectosPiloto.ppt</a> , en donde se requiere el diligenciamiento de una diapositiva con el Cronograma de Hitos del proyecto.	Duración del proyecto: <ul style="list-style-type: none"> <li>• 3 meses: 1</li> <li>• 4 meses: 0.8</li> <li>• Hasta 5 meses: 0.5</li> </ul>

#### 4.4. Criterio 45: Habilidades del equipo en herramientas de Big Data

Este criterio tiene un peso total del 20% para la evaluación del proyecto, el criterio se divide en el subcriterio de Habilidad con servicios de Microsoft - Azure, a continuación, la descripción del subcriterio:

Subcriterio	Descripción	Tipos de Evidencias	Puntajes asignados
Habilidad con servicios de Microsoft - Azure	La plataforma Data Sandbox esta implementada con herramientas y servicios ofrecidos por Microsoft Azure, tener habilidades en el uso de estas tecnologías afianza y agiliza el desarrollo de las iniciativas.	Adjuntar documentación que certifique la experiencia (Nombrar proyectos donde ha participado) en el manejo de servicios y herramientas tecnológicas de Big Data disponibles en la plataforma de Azure. Si a realizado cursos de entrenamiento relacionados con Big Data en la plataforma de Azure añadir los certificados respectivos.	Habilidad con servicios de Microsoft Azure: <p>Integrantes con experiencia en Azure.</p> <ul style="list-style-type: none"> <li>• 7 o más integrantes: 1</li> <li>• 6 integrantes: 0.86</li> <li>• 5 integrantes: 0.71</li> <li>• 4 integrantes: 0.57</li> <li>• 3 integrantes: 0.43</li> <li>• 2 integrantes: 0.29</li> <li>• 1 integrantes: 0.14</li> <li>• 0 integrantes: 0</li> </ul>

En **caso de empate** para acceder a uno o más espacios de los **ocho (8) disponibles** en el Data Sandbox, se **citará a las entidades empatadas a una misma sesión virtual**, en la que deberán **presentar al comité evaluador el proyecto y equipo de trabajo**, en esta misma sesión se determinará a cuál entidad o entidades se les otorgará el beneficio de ingresar al Data Sandbox.

Todos los puntajes serán asignados de acuerdo con las evidencias en la matriz de evaluación del anexo [Criterios de evaluación.xlsx](#)

## 5. Acerca de la generación de capacidades

Se tienen estimadas realizar en total nueve (9) sesiones de formación y capacitación a las entidades y equipos beneficiarios de la convocatoria, los cuales contemplaran algunas de las siguientes temáticas:

Título capacitación	Horas de duración
Características y aplicabilidad de los servicios de Nube Pública	4
Modelos de Servicio: Software como servicio (SaaS) – Plataforma como servicio (PaaS) e Infraestructura como Servicio (IaaS)	4
Características para la adquisición de los servicios de nube pública	4
Configuración de autenticación multifactor	4
Descripción, Características e ingreso a la consola de aprovisionamiento de servicios	4
Creación y configuración de instancias básicas y avanzadas: Máquinas virtuales, bases de datos, almacenamiento, seguridad, redes, contenedores, servicios serverless.	4
Creación y configuración de VPC (Virtual Private Cloud)	4
Características y configuración de reglas para realizar auto escalamiento	4
Modificación de instancias ya creadas	4
Acceso a las instancias	4
Configuración de elementos de red para el servicio de nube pública	4
Creación y configuración de usuarios y roles	4
Configuración de monitoreo básico y avanzado, y generación de reportes.	4
Consumo y facturación de los servicios	4
Configuración y creación de Backups y Snapshots.	4
Configuración de capas de seguridad adicionales para los servicios	4
Funcionalidades básicas y avanzadas de los Productos de nube pública	4
Fallas y problemas comunes presentados en los ambientes de nube pública	4
Solución de problemas que se presenten sobre los servicios	4
Descripción y configuración (en caso de aplicar) en servicios relacionados con IoT, Blockchain, Big Data, Machine Learning.	4

## 6. Acerca de la entrega de los servicios de Nube

La infraestructura que se entregará a las entidades para cada proyecto está soportada en la nube de Microsoft Azure, los servicios se habilitarán dependiendo de la necesidad de cada proyecto, debido a que el Data Sandbox es un espacio para pilotaje, la capacidad de datos soportada para cada proyecto no puede ser superior a 2TB, los servicios serán configurados con capacidades de procesamiento específicas para cada proyecto, previas reuniones con los equipos de trabajo, de igual manera si se evidencia la necesidad de aumentar las capacidades y la plataforma lo permite, éstas serán aumentadas proporcionalmente, una vez terminada el desarrollo de los proyectos y publicados los resultados, las entidades realizaran copia de todos estos insumos, se cerrará el proyecto formalmente y se procederá a eliminar la instancia o espacio asignado a las entidades.

La plataforma soporta servicios como Storage Account, Azure Databricks, Data Factory, Azure Machine Learning, Azure SQL entre otros, si la entidad requiere algún servicio en especial que este dentro del acuerdo marco de la plataforma Data Sandbox, este será habilitado para aprovechar al máximo este espacio colaborativo, es necesario que las entidades acepten la detención de los servicios cuando estos no se estén utilizando.

Tipo Proyectos	Infraestructura	Almacenamiento y procesamiento
Proyectos Grandes Alto consumo mensual	Funcionalidades básicas y avanzadas de los Productos de nube pública, ejemplo servicios cómo Databricks, Apache Hadoop on HDInsight, Apache Spark on HDInsight, Data Factory, entre otros de los servicios del Data Sandbox.	Mayor a 1TB y hasta 2TB
Proyectos medianos consumo medio mensual	Funcionalidades básicas y avanzadas de los Productos de nube pública, ejemplo servicios cómo Databricks, Apache Hadoop on HDInsight, Apache Spark on HDInsight, Data Factory, entre otros de los servicios del Data Sandbox.	Mayor a 500GB y hasta 1TB
Proyectos pequeños	Funcionalidades básicas y avanzadas de los Productos de nube pública, ejemplo servicios cómo Databricks, Apache Hadoop on HDInsight, Apache Spark on HDInsight, Data Factory, entre otros de los servicios del Data Sandbox.	Mayor a 300GB y hasta 500GB

## 7. Contrapartida de las Entidades

Las Entidades que hagan uso del Data Sandbox deberán compartir y publicar algunos de los siguientes resultados obtenidos por el desarrollo de los proyectos piloto en la sección de usos del portal de datos abiertos del estado colombiano - [herramientas.datos.gov.co/es/usos](https://herramientas.datos.gov.co/es/usos), entre los resultados publicables serán:

- **Datos:** Una muestra del universo de datos utilizado en el desarrollo del proyecto piloto que permita la comprobación de los algoritmos compartidos. Estos datos deben cumplir con las leyes

vigentes de protección de datos personales y acceso a información clasificada o reservada descrita en la Ley de Transparencia. Para la publicación es indispensable la anonimización de los datos en los casos en los que aplique.

- **Documentación:** Manuales técnicos, manuales de Usuario, diagramas de arquitectura, diagramas Entidad relación.
- **Algoritmos:** Los algoritmos desarrollados en código abierto o software libre utilizados en los procesos de análisis y/o visualización de los datos.
- **Metodología:** La metodología utilizada en el desarrollo de los proyectos piloto requieren ser documentadas y compartidas para ser utilizadas por otras Entidades.
- Resultados de aprendizaje en **materia de ética de los datos**.
- Resultados de aprendizaje en **materia de ética de los algoritmos** y posibles sesgos que se puedan haber documentado e identificado en los algoritmos utilizados en el desarrollo del proyecto

Los resultados del proyecto piloto deben ser documentados y publicados en herramientas de control de versiones como Git, GitHub, Bitbucket, entre otras opciones que determine la Entidad y adicionalmente se publicará la descripción del proyecto en la sección de usos del portal de datos abiertos del estado colombiano - [herramientas.datos.gov.co/es/usos](http://herramientas.datos.gov.co/es/usos), todo lo anterior con la posibilidad de comprobación y reproducibilidad de los resultados del proyecto por parte de cualquier Usuario que desee reutilizar los algoritmos o demás resultados.

## 8. Cronograma de la convocatoria

La actual convocatoria de proyectos de Big Data presenta las siguientes fechas de ejecución:

Actividad	Descripción	Fechas importantes
Inscripción a la convocatoria	Las entidades públicas podrán inscribir sus proyectos diligenciando el siguiente <a href="#">formulario</a> y adjuntado las evidencias.  Nota: Una vez cargados los soportes en el formulario no se podrán volver a cargar nuevos soportes	<b>Apertura de la convocatoria:</b> 23/03/2021 <b>Cierre de la convocatoria:</b> 31/03/2021
Presentación de proyectos	Las entidades presentarán en un pitch de hasta 15 minutos por cada postulación de proyecto.	<b>Inicio:</b> 05/04/2021 <b>Fin:</b> 09/04/2021
Evaluación de los proyectos	Se evaluarán los proyectos conforme a las evidencias entregadas en el formulario de inscripción de la convocatoria, no se tendrán en cuenta ningún soporte adicional que no haya sido entregado en la inscripción.	<b>Inicio:</b> 24/03/2021 <b>Fin:</b> 09/04/2021
Publicación de resultados	Se publicarán los resultados de la evaluación de los proyectos con sus respectivos puntajes.	12/04/2021

Actividad	Descripción	Fechas importantes
Inicio de proyectos en el Data Sandbox	Se iniciarán las reuniones con las entidades seleccionadas para la entrega del ambiente en el Data Sandbox, adicionalmente se iniciarán los seguimientos al desarrollo de los proyectos.	Semana del 12 al 16 de abril de 2021

## 9. Propiedad intelectual

La propiedad intelectual sobre los resultados derivados del uso de Data Sandbox: El MinTIC se reserva los derechos patrimoniales de la Plataforma de Data Sandbox en su calidad de administrador.

La propiedad intelectual sobre los resultados derivados del uso de Data Sandbox: El titular de los derechos patrimoniales sobre los resultados derivados del uso de los algoritmos, datos, software libre, entre otros productos del proyecto piloto son de titularidad exclusiva de la Entidad. No obstante, La Entidad otorgará una licencia de uso libre, gratuita y permanente para uso del MinTIC, de las Entidades públicas participantes, emprendedores, entre otros actores que deseen reutilizar los resultados derivados del proyecto piloto. La Entidad o cualquier Usuario o participante deberá citar en todas sus publicaciones lo siguiente: “Los resultados del proyecto piloto en el Data Sandbox fue realizado con el apoyo del Ministerio de Tecnologías de la Información y las Comunicaciones”.

Los nombres comerciales, marcas o signos distintivos que aparecen o a los que se hace alusión al Data Sandbox o en la Plataforma de Datos Abiertos del Estado Colombiano, entre otros relacionados, pertenecen al MinTIC o a sus debidos propietarios y se encuentran protegidos por la legislación vigente al respecto.

Todos los logotipos y marcas de la Plataforma son de propiedad de MinTIC y su uso ha sido autorizado por sus titulares a MinTIC, siendo, en todo caso, los titulares los responsables de cualquier posible controversia que pudiera darse respecto de ellos. Los titulares de dichas marcas y logotipos se reservan el derecho de entablar las acciones legales que consideren convenientes para hacer valer sus derechos tanto en Colombia como en el exterior.

En caso de reclamaciones que pudieran interponerse por los Usuarios o por terceros en relación con posibles incumplimientos de los derechos de propiedad intelectual sobre cualquiera de los contenidos de los sitios descritos anteriormente, deberán dirigirse a la siguiente dirección de correo electrónico: [usoweb@mintic.gov.co](mailto:usoweb@mintic.gov.co), una vez notificado a este correo, dicho contenido será automáticamente eliminado de los sitios hasta que quien haya publicado el contenido en disputa haya resuelto el conflicto con quien envía la reclamación.

## 10. Modificaciones a los términos de referencia

El MINTIC se reservará el derecho de modificar los términos y condiciones de uso aquí contenidas a su libre elección y en cualquier momento, cualquier cambio será informado y publicado oportunamente en la página web. El Usuario se compromete a revisar periódicamente esta sección para estar informado de

tales modificaciones y cada nuevo acceso del usuario a la página será considerado una aceptación tácita de las nuevas condiciones.

## 11. Anexos

### **Anexo No. 1. Criterios Evaluación de Proyectos**

Contiene los criterios de evaluación a ser tenidos en cuenta para la selección de los proyectos a ingresar al Data Sandbox.

[https://herramientas.datos.gov.co/sites/default/files/Criterios%20de%20evaluaci%C3%B3n\\_0.xlsx](https://herramientas.datos.gov.co/sites/default/files/Criterios%20de%20evaluaci%C3%B3n_0.xlsx)

### **Anexo No. 2. Términos y condiciones de uso del Data Sandbox**

<https://sandbox.datos.gov.co/#!/terminosCondiciones>