
Informe Final de Auditoría

Con Informe Ejecutivo

Proyecto N° 10.21.04

BUENOS AIRES DATA

Auditoría de Sistemas

Período 2020

Buenos Aires, Septiembre 2022



AUDITORIA GENERAL
DE LA CIUDAD DE BUENOS AIRES

“Año del 40° Aniversario de la Guerra de Malvinas. En homenaje a los veteranos
y caídos en la defensa de las Islas Malvinas y el Atlántico Sur”

AUDITORIA GENERAL DE LA CIUDAD DE BUENOS AIRES

JEAN JAURES 220 - CIUDAD DE BUENOS AIRES

Presidente

Dra. Mariana Inés GAGLIARDI

Auditores Generales

Dr. Juan José CALANDRI

Dr. Pablo CLUSELLAS

Cdra. Mariela Giselle COLETTA

Lic. María Raquel HERRERO

Dr. Daniel Agustín PRESTI

Dr. Lisandro Mariano TESZKIEWICZ

Código de Proyecto: 10.21.04

Nombre del Proyecto Auditoría: Buenos Aires Data.

Tipo de auditoría: Sistemas.

Organismo auditado: Dirección General de Calidad Institucional y Gobierno Abierto (DGCIGA). Dirección General de Arquitectura de Datos (DGARD).

Objeto: Sistema Buenos Aires Data.

Objetivo: Evaluar el entorno informático, tecnológico y operativo del Sistema.

Alcance: Analizar la continuidad, seguridad física y lógica de los sistemas, los procesos involucrados y manejo de la información.

Jurisdicción / Unidad Ejecutora: Jurisdicción 20 Jefatura de Gobierno - Unidad Ejecutora 2003 D.G. Calidad Institucional y Gobierno Abierto

Presupuesto: Jurisdicción 20 - Unidad Ejecutora 2003 - Programa 54 “Planificación de la Calidad Institucional” / Jurisdicción 21 - Unidad Ejecutora 7311 – Programa 109 “Ciencias de la Información”.

Período a Auditar: 2020

Directora de Proyecto: Jimena Dominguez

Supervisor: Abg. Mariano Monzón Font

Las tareas de auditoría comenzaron en marzo de 2021.

FECHA DE APROBACION DEL INFORME: 14 DE SEPTIEMBRE DE 2022

APROBADO POR: UNANIMIDAD

INFORME EJECUTIVO

Lugar y fecha de emisión	Ciudad Autónoma de Buenos Aires, 14 de septiembre 2022.
Código de Proyecto	10.21.04
Denominación del Proyecto	Buenos Aires Data.
Tipo de Auditoría	Sistemas.
Dirección General	Dirección General de Sistemas de Información.
Período bajo examen	2020
Objeto de la Auditoría	Buenos Aires Data.
Objetivo de la Auditoría	Evaluar el entorno informático, tecnológico y operativo del Sistema.
Alcance	Analizar la continuidad, seguridad física y lógica de los sistemas, los procesos involucrados y manejo de la información.
Limitaciones al Alcance	No existieron limitaciones al alcance.
Observaciones Relevantes	<ul style="list-style-type: none"> ➤ Alinear, Planificar y Organizar - APO02 <ul style="list-style-type: none"> • Gestionar la estrategia <ol style="list-style-type: none"> 1. La información que generan las reparticiones del GCABA se halla distribuida en cada una de ellas, no se encuentra centralizada, vinculada o compartida salvo para casos y sistemas específicos. No existe una política general para integrar la información generada por las reparticiones del GCABA. Así, para publicar en el portal de datos públicos abiertos Buenos Aires Data la DGARD y DGCIGA necesitan caso por caso solicitar a cada repartición que comparta los datos que genera.

La definición de una estrategia de gestión integral de la información pública para el GCABA es de suma importancia no solo para el objetivo de un gobierno abierto sino para la toma de decisiones de políticas públicas con base en la evidencia.

➤ Alinear, Planificar y Organizar - APO03

• Gestionar la arquitectura empresarial

2. En principio, cabe señalar que la definición es compleja en cuanto a qué criterios estratégicos deben tomarse para la publicación información generada por las reparticiones del GCABA; y que las áreas intervinientes DGCIGA y DGARD buscan activamente la apertura de datos en poder de otras reparticiones que son demandados por actores de la sociedad civil, academia y prensa.

No obstante, no se encuentra definido el universo de la información pública que generan las distintas reparticiones del GCABA y que luego serían publicadas en el portal de datos abiertos Buenos Aires Data. En su defecto, tampoco existen criterios o definiciones estratégicos de qué información es más relevante para ser publicada en el portal. Definir el universo o criterios de importancia sobre la información a publicar en el portal permite confirmar que la información es la más relevante.

3. La información que va a ser publicada debe ser solicitada para cada caso. El flujo de la información no es continuo ni automático y ésta se ve afectada por desactualizaciones. Las áreas responsables de la publicación de la información (DGARD – DGCIGA) se ven afectadas por el accionar de las áreas productoras de la información cuando incumplen con la actualización de los datos. Solo pueden indicarle al área que debe cumplir con la frecuencia de actualización, y no pueden garantizar que la información se encuentre actualizada. Asimismo, mucha de la información es de actualización eventual. Contar con información actualizada es uno de los pilares en los que debe sostenerse un portal de gobierno abierto.

4. Existen pocas soluciones (APIs) que permitan la disponibilidad inmediata de la información a medida que es generada. Una mayor disponibilidad de la información redundaría en una mayor calidad de las prestaciones del sistema.

➤ Alinear, Planificar y Organizar - APO07

- Gestionar los recursos humanos

5. Se relevó que la gran mayoría de los agentes que prestan tareas técnicas dentro de la DGARD se encuentran contratados bajo la modalidad de Locación de Obra y Servicios, no perteneciendo a la dotación de agentes del GCABA. La falta de formalización en la relación laboral de los agentes podría generar inestabilidad en la funcionalidad del sistema por la rotación del recurso humano en las tareas asignadas.

➤ Alinear, Planificar y Organizar - APO14

- Gestionar los datos

6. Se halló que gran cantidad de los datasets tenían el plazo de la frecuencia de actualización vencido. Asimismo, que los datos contenidos dentro de algunos de los datasets se encontraban desactualizados, con errores y omisiones en la información contenida.

7. No existen controles automáticos por medio de softwares que intenten garantizar que se cumplan los plazos de actualización por parte de las reparticiones, y que la información se encuentre efectivamente actualizada, se encuentre normalizada y no contenga errores o esté incompleta.

<p>Conclusión / Dictamen</p>	<p>Como primer punto a abordar, es necesario subrayar el importante objetivo de gobierno abierto buscado por el portal y las prestaciones que provee. El Gobierno abierto implica tanto la apertura de los datos públicos como de procesos internos de la administración, basándose en los principios de transparencia, colaboración y participación.</p> <p>En su gestión diaria, El GCABA genera muchísima información en cada una de sus reparticiones. Pero dicha información suele quedar circunscripta a la repartición, no se halla disponible para ser accedida, compartida o consumida.</p> <p>Si bien cada repartición es “dueña” del dato (<i>data owner</i>), eso no significa que deba utilizarlo -si es que lo hace- para sus fines organizativos propios únicamente. No es correcto que ellas puedan considerar a su información como privativa y vuelvan celosas guardianas de la misma.</p> <p>En el presente informe se pudo comprobar que la información se encuentra localizada en cada repartición, y para accederla es necesario lograr acuerdos de apertura de datos entre la repartición y las áreas encargadas de la publicación de la información pública en el portal de datos abiertos del GCABA. También se encontró que el flujo de la información no es continuo ni automático y ésta se ve afectada por desactualizaciones, problemas en la normalización y formatos de la información.</p> <p>En cuanto a la gestión de la información, el GCABA debe considerarse a sí mismo como una sola organización, y debe tratar la información de una manera íntegra, consistente, interrelacionada, compatible y coherente.</p> <p>No debe entenderse lo anterior que se propone un área que sea la única que se encargue de recabar la información de manera centralizada. Sino que se indica como necesario definir políticas y articular soluciones tecnológicas para que la información generada y tratada por una repartición sea accesible (en las condiciones de seguridad y privacidad que el caso demande) y utilizable por otras, y en especial la conducción del poder ejecutivo.</p> <p>La información producida por un sector del gobierno seguramente sea un insumo de otro y, de lleno, un punto de</p>
---	--

“Año del 40° Aniversario de la Guerra de Malvinas. En homenaje a los veteranos y caídos en la defensa de las Islas Malvinas y el Atlántico Sur”

	<p>partida para las decisiones en la gestión que realizan los altos funcionarios.</p> <p>Todo lo anterior redundaría en mejorar la experiencia de los ciudadanos y el público en general, ayudar a definir las políticas públicas en base a la evidencia, y lograr la economía, eficacia, eficiencia, equidad y sostenibilidad en las políticas y acciones de gobierno.</p>
Palabras Claves	Gobierno Abierto, BA Data, datos abiertos, datos públicos, gestión de la información

**INFORME FINAL DE AUDITORÍA
“BUENOS AIRES DATA”
PROYECTO N° 10.21.04**

DESTINATARIO

Vicepresidente 1°
Legislatura de la Ciudad Autónoma de Buenos Aires
Lic. Emmanuel Ferrario
S / D

En uso de las facultades conferidas por los artículos 131, 132 y 136 de la Ley 70 de la Ciudad Autónoma de Buenos Aires, y conforme a lo dispuesto en el artículo 135 de la Constitución de la Ciudad, la Auditoría General de la Ciudad de Buenos Aires ha procedido a efectuar una auditoría informática en el ámbito de las reparticiones Dirección General de Calidad Institucional y Gobierno Abierto (DGCIGA) y Dirección General de Arquitectura de Datos (DGARD) con el objeto detallado en el apartado siguiente.

I. OBJETO DE AUDITORIA

Sistema Buenos Aires Data.

II. OBJETIVO DE LA AUDITORIA

Evaluar el entorno informático, tecnológico y operativo del Sistema.

III. ALCANCE DEL EXAMEN

Analizar la continuidad, seguridad física y lógica de los sistemas, los procesos involucrados y manejo de la información.

III.1.- Marco Normativo de la Auditoría de Sistemas:

Se utilizó el siguiente marco normativo:

Normas Básicas de Auditoría Externa¹ de la AGCBA y las Normas Básicas de Auditoría de Sistemas² de la AGCBA, la Ley 70 CABA, Ley 325 CABA y complementarias.

Normas COBIT 2019³ (Control Objectives for Information and related Technology) de la Information Systems Audit and Control Association (ISACA). Capítulo Buenos Aires.

Las normas y recomendaciones de Tecnologías de la Información (TI) establecidas por la ASINF (Res 177/ASInf/13⁴ y ampliaciones).

III.2.- Procedimientos realizados.

Se remitió nota al Director General de la Dirección General de Calidad Institucional y Gobierno Abierto (DGCIGA).

Se efectuó una reunión con el Director General de Calidad Institucional y Gobierno Abierto, se presentó al equipo de trabajo, se explicó el alcance del proyecto, se aclararon temas relacionados al mismo y se acordaron los pasos a seguir.

Se envió una nota de pedido de información a la Dirección General de Calidad Institucional y Gobierno Abierto (DGCIGA), inquirendo sobre su normativa, presupuesto, personal, funcionamiento, procesos, sistemas informáticos y tecnologías.

Se efectuó una reunión con la Directora General de Arquitectura de Datos (DGARD), se presentó al equipo de trabajo, se explicó el alcance del proyecto, se aclararon temas relacionados al mismo y se acordaron los pasos a seguir.

Se remitió una nota de pedido de información a la Dirección General de Arquitectura de Datos (DGARD), inquirendo sobre su normativa, presupuesto, personal, funcionamiento, procesos, sistemas informáticos y tecnologías.

¹ Las Normas Básicas de Auditoría Externa de la AGCBA se encuentran publicadas es https://www.agcba.gov.ar/web/doc/ni-normas_basicas_de_auditoria_externa.pdf [Accedido el 9/10/21]

² Normas Básicas de Auditoría de Sistemas. Se encuentran en https://www.agcba.gov.ar/web/doc/ni-normas_basicas_de_auditoria_sistemas.pdf [Accedido el 9/10/21]

³ Normas COBIT 2019. <https://www.isaca.org/resources/cobit> [Accedido el 9/10/21]

⁴ Resolución N°177/ASInf/13 https://www.buenosaires.gob.ar/sites/gcaba/files/rs-2013-177--asinf_4.pdf [Accedidos el 9/10/21]

Se creó una carpeta compartida mediante la solución en la nube de Onedrive, otorgándole permisos de edición a los emails de los funcionarios y agentes del órgano dedicados a esta auditoría.

Se realizaron entrevistas virtuales con la Gerenta Operativa de Gobierno Abierto de la Dirección General de Calidad Institucional y Gobierno Abierto (DGCIGA) y el Líder de Proyecto de BA Data de la Dirección General de Arquitectura de Datos (DGARD).

Se revisaron en el Presupuesto 2020 y en la Cuenta de Inversión 2020 los programas: Programa 54 - Planificación de la Calidad Institucional - Jurisdicción 20 Jefatura de Gobierno - Unidad Ejecutora 2003 D.G. Calidad Institucional y Gobierno Abierto; Programa 109 “Ciencias de la Información”, Unidad Ejecutora 7311 Subsecretaría Políticas Públicas basadas en Evidencias (SSPPBE), Jurisdicción 21 Jefatura de Gabinete de Ministros.

Se efectuó un análisis de los datasets publicados en el portal sobre un muestreo representativo respecto de su actualización e integridad. Se comprobó la funcionalidad de las APIs publicadas.

IV.- Limitaciones Al Alcance.

No existieron limitaciones al alcance.

V.- ACLARACIONES PREVIAS

1.- Buenos Aires Data⁵ - Introducción:

Los datos se han convertido en el insumo principal de la economía, y se transformaron en el activo más valioso y estratégico con el que cuentan las organizaciones. Con el surgimiento de las nuevas tecnologías (Inteligencia Artificial, Big Data, Data Mining), las políticas de datos abiertos se han vuelto un ingrediente más para promover el desarrollo y crecimiento económico al disponer de un recurso fundamental para la toma de decisiones.

Buenos Aires Data es un portal de datos abiertos creado en el año 2012, en la plataforma se recopilan más de 400 conjuntos de datos (*datasets*⁶) de

⁵ <https://www.buenosaires.gob.ar/datosabiertos> o <https://data.buenosaires.gob.ar/dataset/>
[Accedido el 03/02/2022]

⁶ Un dataset es una colección ordenada de datos. Mientras se manejan los datos, el conjunto de datos puede ser un montón de tablas, esquemas y otros objetos. Los datos están esencialmente organizados en un determinado modelo que ayuda a procesar la información necesaria.

diferentes áreas de Gobierno de la Ciudad Autónoma de Buenos Aires (GCABA). El objetivo es facilitar el acceso, descubrimiento y reutilización de los datos recolectados dentro de una política de gobierno abierto⁷.

Los datos se encuentran publicados en formato abierto para que cualquier persona los pueda descargar y reutilizar para crear nuevos análisis, aplicaciones o ideas.

De esta manera los ciudadanos y empresas pueden aprovechar y crear productos y servicios de valor agregado que permitan mejorar el modo en que los habitantes se muevan por la ciudad, accedan a la vivienda y hagan turismo, etc.

2.- Reparticiones / Áreas intervinientes.

El portal “Buenos Aires Data” es gestionado en conjunto por 2 áreas distintas del GCABA: La Dirección General de Calidad Institucional y Gobierno Abierto (DGCIGA), dentro de la Secretaría General y Relaciones Internacionales (SGYRI), que se encarga de la atención y relaciones con el público; y la Dirección General de Arquitectura de Datos (DGARD), dentro de la Jefatura de Gabinete de Ministros (MJGGC), que aporta los aspectos técnicos sobre los *dat@sets* y su publicación.

En la redacción del presente informe se utilizarán como sinónimos las palabras “áreas” y “reparticiones”.

2.1.- Dirección General de Calidad Institucional y Gobierno Abierto (DGCIGA):

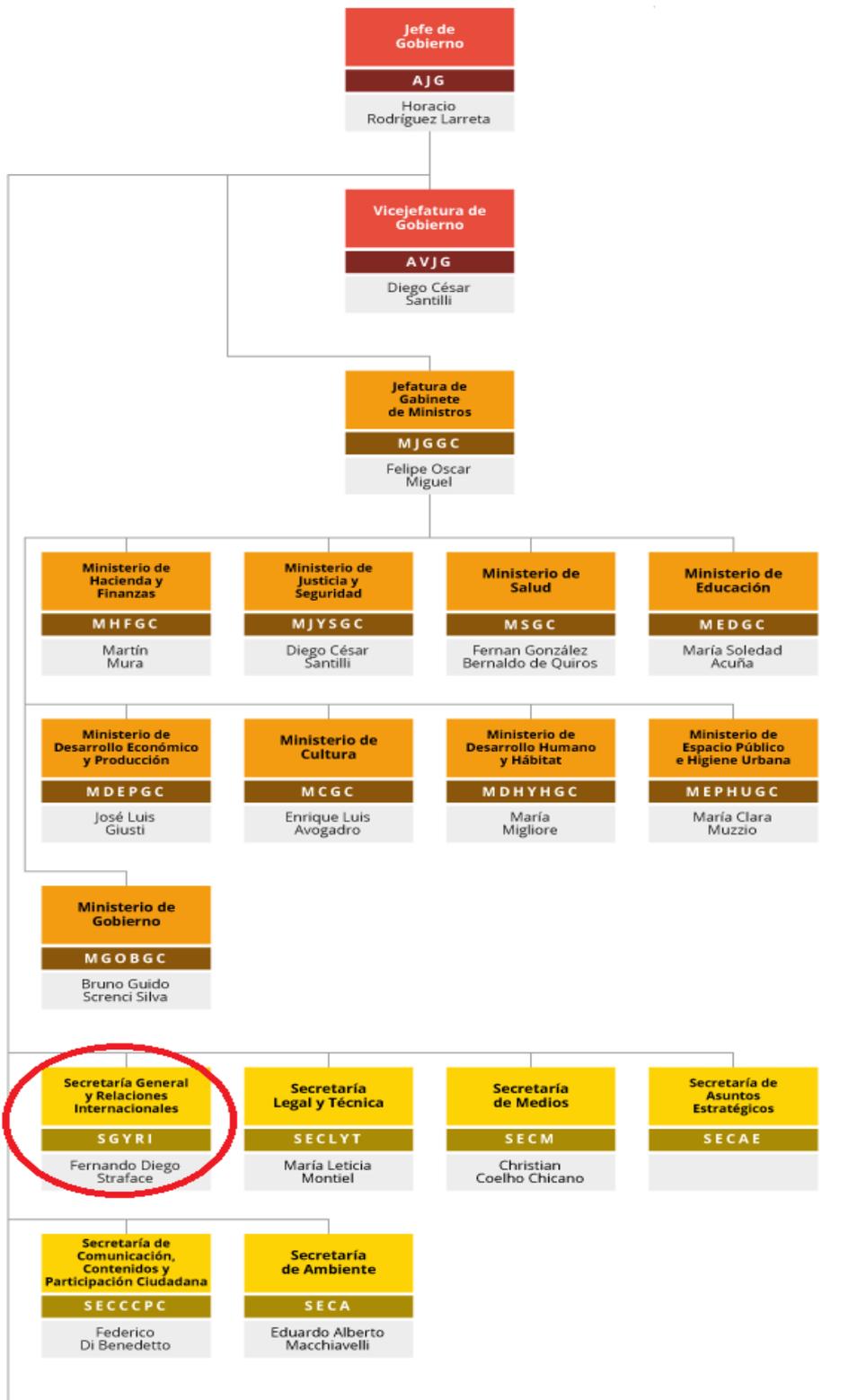
2.1.1- Normativa - Estructura Orgánico-Funcional.

La Dirección General de Calidad Institucional y Gobierno Abierto (DGCIGA) depende de la Subsecretaría de Gobierno Abierto y Rendición de Cuentas (SSGOARCU), que a su vez se encuentra jerárquicamente bajo la Secretaría General y Relaciones Internacionales (SGYRI), y ésta subordinada al Jefe de Gobierno.

⁷ https://es.wikipedia.org/wiki/Gobierno_abierto [Accedido el 03/02/2022]

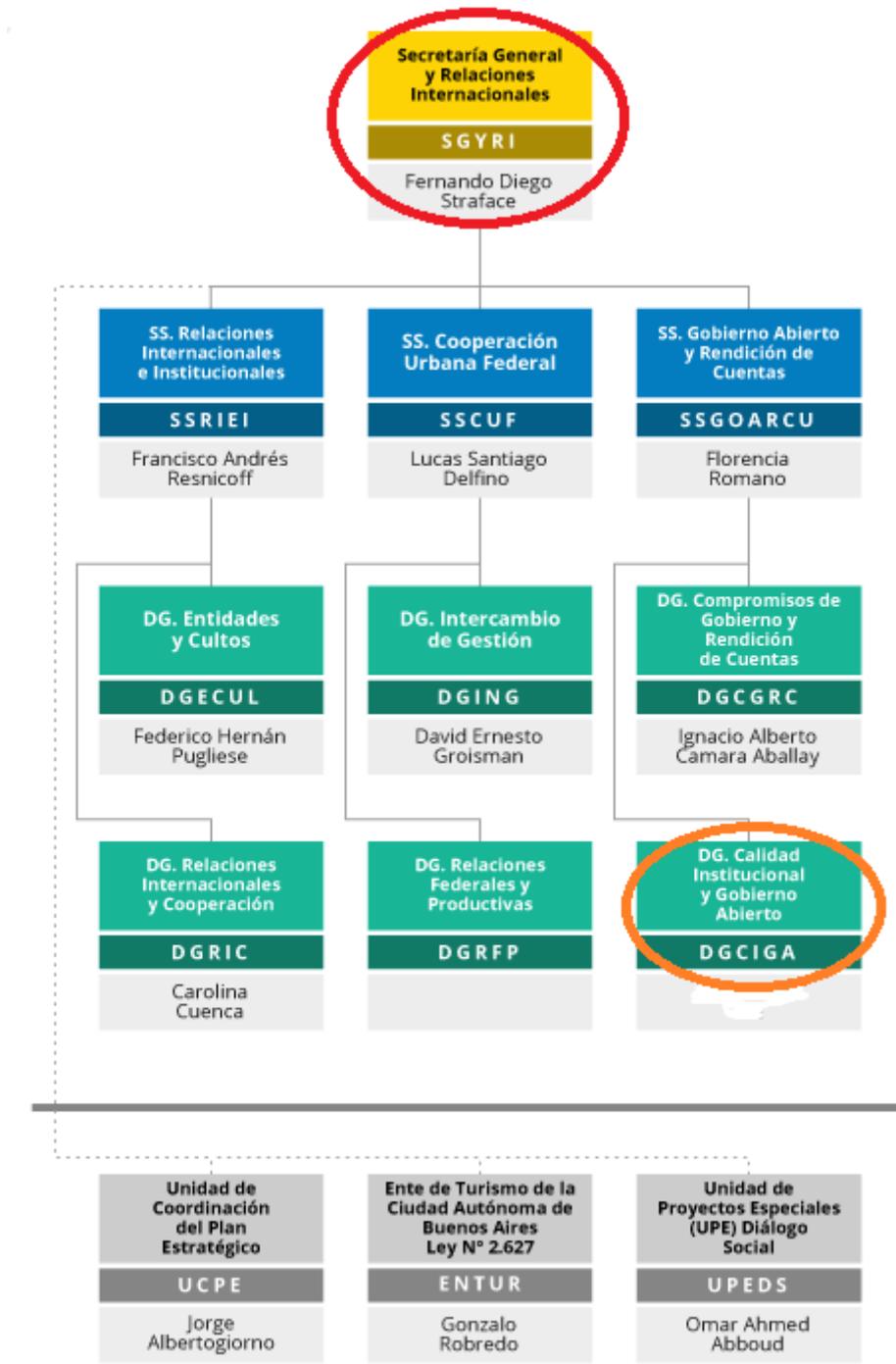
“Año del 40° Aniversario de la Guerra de Malvinas. En homenaje a los veteranos y caídos en la defensa de las Islas Malvinas y el Atlántico Sur”

organigrama



“Año del 40° Aniversario de la Guerra de Malvinas. En homenaje a los veteranos y caídos en la defensa de las Islas Malvinas y el Atlántico Sur”

organigrama



2.1.2.- Competencia - Responsabilidades Primarias.

Las responsabilidades primarias fueron definidas por el Decreto GCABA 116/2021⁸, que modifica al anterior Decreto GCABA 463/2019⁹, y se enumeran en el Anexo I.

Entre las más relevantes, según el objeto de esta auditoría, podemos destacar:

- Planificar y ejecutar las acciones de apertura de datos mediante el desarrollo de bases de datos de libre acceso, en coordinación con las áreas competentes y promoviendo la reutilización de datos por parte de la sociedad civil y el sector privado.
- Administrar plataformas de datos abiertos de la Ciudad Autónoma de Buenos Aires, incluyendo BA Data y BA Obras, en coordinación con la Secretaría de Innovación y Transformación Digital de la Jefatura de Gabinete de Ministros.

2.1.3.- Presupuesto – Jurisdicción – Programa.

El presupuesto financiero sancionado para el Programa 54 “Planificación de la Calidad Institucional”, Unidad Ejecutora 2003 DG Calidad Institucional y Gobierno Abierto (DGCIGA), Jurisdicción 20 Jefatura de Gobierno; se expone a continuación:

⁸ <https://boletinoficial.buenosaires.gob.ar/normativaba/norma/547755> [Accedido el 04/02/21]

⁹ <https://boletinoficial.buenosaires.gob.ar/normativaba/norma/500340> [Accedido el 04/02/21]

“Año del 40° Aniversario de la Guerra de Malvinas. En homenaje a los veteranos y caídos en la defensa de las Islas Malvinas y el Atlántico Sur”

Programa: 54 PLANIFICACIÓN DE LA CALIDAD INSTITUCIONAL

Unidad Ejecutora: DIRECCION GENERAL CALIDAD INSTITUCIONAL Y GOBIERNO ABIERTO

Jurisdicción: 20.JEFATURA DE GOBIERNO

Finalidad: Administración Gubernamental

Función: Dirección ejecutiva

PRESUPUESTO FINANCIERO	
Inciso	
Principal	IMPORTE
Gastos en personal	9.813.235
Personal Permanente	7.306.427
Asignaciones familiares	69
Asistencia social al personal	146.597
Gabinete de autoridades superiores	2.360.142
Bienes de consumo	194.000
Productos alimenticios, agropecuarios y forestales	150.000
Pulpa,papel, cartón y sus productos	16.500
Otros bienes de consumo	27.500
Servicios no personales	4.483.657
Mantenimiento, reparación y limpieza	330.000
Servicios profesionales, técnicos y operativos	2.310.819
Servicios Especializados, Comerciales y Financieros	881.937
Pasajes, viáticos y movilidad	354.101
Otros servicios	606.800
TOTAL	14.490.892

En el Anexo II constan los objetivos previstos para el Programa 54 en la Ley de Presupuesto 2020.

Dos de ellos son dignos de ser destacados por su vinculación con el objeto de la presente auditoría:

- Coordinación de la Política de Datos Abierto del GCBA.
- Mantener y actualizar los portales y repositorios Github correspondientes a las iniciativas BA Data y BA Obras.

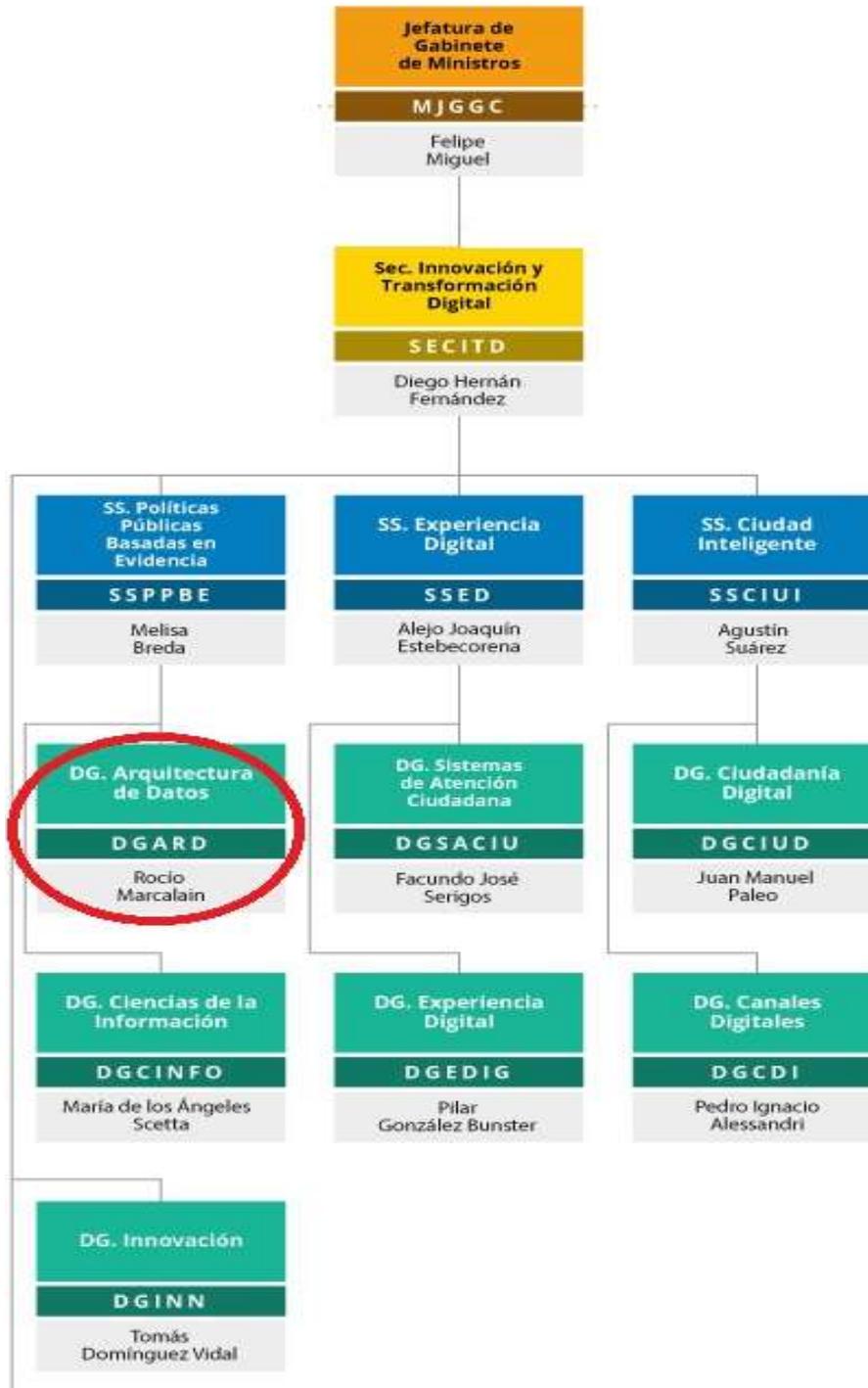
2.2. – Dirección General de Arquitectura de Datos.

2.2.1- Normativa - Estructura Orgánico-Funcional.

La Dirección General de Arquitectura de Datos (DGARD) depende de la Subsecretaría de Políticas Públicas Basada en Evidencia (SSPPBE), que a su vez se encuentra jerárquicamente bajo la Secretaría de Innovación y

“Año del 40° Aniversario de la Guerra de Malvinas. En homenaje a los veteranos y caídos en la defensa de las Islas Malvinas y el Atlántico Sur”

Transformación Digital (SECITD), y ésta subordinada a la Jefatura de Gabinete de Ministros (MJGGC).



2.2.2.- Responsabilidades Primarias

Las Responsabilidades primarias fueron definidas por el Decreto GCABA Nº282/2021¹⁰ el cual modificó la estructura organizativa de la Jefatura de Gabinete de Ministros aprobada por el Decreto GCABA Nº463/2019, ya nombrado.

Entre ellas se destacan las vinculadas con el objeto de esta auditoría:

- Mantener y administrar el portal de datos de la Ciudad Autónoma de Buenos Aires, en coordinación con la Secretaría General y Relaciones Internacionales. Realizar el procesamiento y control de los datasets, disponibilizándolos para su publicación en el portal de datos abiertos, y proveer asesoramiento a las áreas del Gobierno de la Ciudad Autónoma de Buenos Aires para mejorar la calidad de la información, en coordinación con la Secretaría General y Relaciones Internacionales. Implementar y administrar la infraestructura de datos espaciales de la Ciudad Autónoma de Buenos Aires.

En el Anexo III se enuncian las responsabilidades primarias de la DGARD, así como también las acciones de las áreas que de ella dependen.

2.2.3. - Presupuesto – Jurisdicción – Programa.

El presupuesto financiero de sanción para el Programa 109 “Ciencias de la Información”, Unidad Ejecutora 7311 Subsecretaría Políticas Públicas basadas en Evidencias (SSPPBE), Jurisdicción 21 Jefatura de Gabinete de Ministros; se expone seguidamente:

¹⁰ ² <https://boletinoficial.buenosaires.gob.ar/normativaba/norma/563711> [Accedido el 03/02/2022]

“Año del 40° Aniversario de la Guerra de Malvinas. En homenaje a los veteranos y caídos en la defensa de las Islas Malvinas y el Atlántico Sur”

Programa: 109 CIENCIAS DE LA INFORMACION

Unidad Ejecutora: SUBS.POL.PÚBL.BÁS.EVIDENCIAS
Jurisdicción: 21.JEFATURA DE GABINETE DE MINISTROS
Finalidad: Servicios Sociales
Función: Educación

PRESUPUESTO FINANCIERO	
Inciso	
Principal	IMPORTE
Gastos en personal	19.195.723
Personal Permanente	15.967.708
Personal Transitorio	466.646
Asignaciones familiares	15.355
Asistencia social al personal	286.227
Gabinete de autoridades superiores	2.459.787
Bienes de consumo	285.000
Productos alimenticios, agropecuarios y forestales	120.000
Pulpa,papel, cartón y sus productos	95.000
Otros bienes de consumo	70.000
Servicios no personales	2.559.500
Mantenimiento, reparación y limpieza	62.000
Servicios Especializados, Comerciales y Financieros	2.375.500
Otros servicios	122.000
TOTAL	22.040.223

3.- Personal - Recursos Humanos.

En la DGCIGA son 3 los agentes que trabajan en el proyecto BA Data, de manera parcial, estimando dentro de la carga de la totalidad de los proyectos un 60%. La DGCIGA cuenta con 4 agentes más, los cuales no brindan tareas relacionadas con BA Data.

La DGARD cuenta con 6 agentes que prestan tareas relacionadas con BA Data, estimando en 15% la carga dentro de la totalidad de los proyectos a cargo de la Dirección General. Los perfiles de los agentes son los siguientes:

“Año del 40° Aniversario de la Guerra de Malvinas. En homenaje a los veteranos y caídos en la defensa de las Islas Malvinas y el Atlántico Sur”

Tareas	Perfil técnico	Título	Situación de Revista
Responsable de la gerencia que desarrolla y mantiene el portal BA Data	Especialista en bases de datos y gestión de equipos de Desarrollo	Técnico en electrónica	Gerente Operativo (GOID)
Coordinador de BA Data	Capacitado en datos abiertos, ciencia de datos y gestión de productos de software	Sociólogo	Adscripto (desde otro organismo)
Actualización de datasets en BA Data y brinda asistencia técnica	Programador	Ingeniero en Sistemas	Locación de Obra y Servicios (LOyS)
Desarrollo de software	Programadora	Ingeniera en Petróleo	Locación de Obra y Servicios (LOyS)
Desarrollo de software. Encargado del software Plataforma de Gestión de Datos (PGD)	Programador	Técnico Universitario en Web	Locación de Obra y Servicios (LOyS)
Desarrollo de software	Programador front-end	Estudiante Licenciatura en Sistemas de la Información	Locación de Obra y Servicios (LOyS)

4.- Informes de Auditoría previos relevantes.

Al momento de la confección del presente informe no se han realizado auditorías -externas o internas- sobre el portal BA Data, o sobre las reparticiones Dirección General de Calidad Institucional y Gobierno Abierto (DGCIGA) o Dirección General de Arquitectura de Datos (DGARD).

5.- Servicios recibidos y brindados.

5.1.- Servicios recibidos.

5.1.1- Servicios recibidos de la Agencia de Sistemas de Información (ASINF):

En los servicios recibidos se destacan los referidos a la conectividad, gestión de la infraestructura, políticas de respaldo, contingencia, y mesa de ayuda, que son provistos por la Agencia de Sistemas de Información (ASINF).

5.1.2- Servicios recibidos de proveedores externos:

Durante el año 2020, para el portal del BA Data, a través de la DGCIGA, se realizó la contratación a la empresa Devartis S.A, de los siguientes servicios, en general referidos a la sección “Historia con Datos”:

- Soporte tecnológico para mantenimiento y desarrollo de prestaciones para el portal. Fecha de contratación: Julio de 2020. Monto: \$80.000.

- Rediseño y desarrollo de contenido que permita renovar la página de inicio de la sección "Historias con Datos" incluyendo el almacenamiento para documentar y publicar las visualizaciones, gráficas e historias con datos desarrolladas. También, un repositorio de las diversas historias cuyo objetivo es ofrecer un listado con la identificación de las mismas el cual permita desplegar el detalle de cada historia mediante una ficha técnica. Fecha de contratación: Noviembre de 2020. Monto: \$300.000.

- Servicio de soporte y mantenimiento de la sección “Historias con Datos”, incluyendo los siguientes recursos humanos: 1 diseñador web especialista en UX/UI, 1 desarrollador front end, 1 desarrollador back end y 1 Project manager. Fecha de la contratación: Diciembre de 2020. Monto: \$260.000.

5.1.- Servicios brindados.

Entre las 2 áreas involucradas, ambas se prestan servicios mutuamente conforme se explicita más adelante en el punto “6.2 Tareas que realizan las áreas”.

En cuanto a los servicios brindados, se pueden enumerar los siguientes:

Servicios brindados por la DGCIGA:

Relaciones con la comunidad: Se encarga de establecer un vínculo con la comunidad exterior al gobierno. Recibe demandas y consultas del público, tales como periodistas, organizaciones, ciudadanos, de la comunidad académica y de la comunidad de datos. Así también asiste y soluciona problemas no técnicos del portal.

Relaciones con las reparticiones del GCABA: en asistencia a la DGARD, se vincula con otras reparticiones del GCABA para obtener accesos a datasets de dichas reparticiones y poder publicarlos.

Otros proyectos: la DGCIGA cuenta con otros proyectos como BA Obras, presupuesto abierto, BA Cambio climático, entre otros. Estas plataformas a su vez generan datasets que son publicados en BA Data.

Servicios brindados por la DGARD:

Administra técnicamente el Portal BA Data.

Se vincula y asesora técnicamente a las reparticiones que van a brindar los datos a publicar. Recibe de los datasets, normaliza, integra y controla la información recibida y la carga en el repositorio correspondiente para su publicación.

Soluciona problemas de soporte técnico sobre funcionalidades y datos publicados en el portal.

6.- Información del/os Sistema(s).

6.1.- Portal Buenos Aires Data:

El *home* de BA Data está en <https://www.buenosaires.gob.ar/datosabiertos>. Allí también están alojadas las secciones de Historias y APIs.

Es administrado por la Dirección General Experiencia Digital (DGEDIG).

Se utiliza el sistema de gestión de contenidos Drupal¹¹ para administrar la información.

Desde el home se accede al portal BA Data en <https://data.buenosaires.gob.ar/dataset/>. Allí se encuentran *datasets* que pueden ser descargados y referencias sobre los mismos: frecuencia de actualización, organizaciones (reparticiones) que brindan los datos, temas con los que se vinculan, etc. Es administrado por la Dirección General Arquitectura de Datos.

El portal ha recibido cambios y mejoras durante los años 2020 y 2021. En cuanto a la estética, no sufrió grandes modificaciones; es la misma de las otras páginas de gobierno y es establecida por la Subsecretaría Experiencia Digital (SSED).

6.2- Tareas que realizan las áreas:

La responsabilidad de las tareas sobre el Sistema de BA DATA del GCABA se distribuyen entre la DGARD y DGCIGA de la siguiente manera:

Tarea	Área Responsable
Recepción, control, normalización de los datasets recibidos de las reparticiones.	DGARD
Propuesta de modificación de la plataforma (tecnológica, interfaz, textos) consensuados.	DGARD - DGCIGA
Soportar y Administrar el portal BA Data.	DGARD
Mantener actualizados los datos publicados en el portal, así como incorporar nuevos activos.	DGARD - DGCIGA
Designar nuevos enlaces de datos en las áreas quienes serán los nexos para continuar la política de datos abiertos al interior del GCBA.	DGARD - DGCIGA
Proponer proyectos de apertura para la publicación de nuevos datos.	DGARD
Atención de contactos de mail de los usuarios del portal. Generación de Tableros de Estadística.	DGARD
Planificar e implementar acciones de difusión del portal de datos abiertos hacia la comunidad de datos abiertos externa al GCBA.	DGCIGA

¹¹ Drupal es un sistema de gestión de contenidos o CMS (por sus siglas en inglés, Content Management System) libre (open source), modular y multipropósito que permite publicar en la web artículos, imágenes, archivos, blogs, etc. <https://www.drupal.org/> [Accedido el 03/02/2022]

Cualquier acción de difusión del portal deberá ser consensuada y coordinada entre las partes previamente.	DGARD - DGCIGA
Contacto con usuarios externos (personas, instituciones, empresas).	DGARD - DGCIGA
Desarrollo de Componentes Gráficos de Datos para páginas Web.	DGARD
Administración de Componentes Gráficos de Datos para páginas Web.	DGARD - DGCIGA

Para llevar adelante la coordinación de tareas, las dos áreas tienen reuniones de seguimiento periódicas y trabajan en conjunto para coordinar la agenda de apertura de datos y los proyectos relacionados.

Además, se utiliza un sistema de peticiones para elevar las solicitudes desde la DGCIGA hacia la Dirección General de Arquitectura de Datos (DGARD) a través de tickets en el sistema Redmine¹².

6.3 - Proyectos de desarrollo.

Entre los planes de corto, mediano y largo plazo relacionados con el Sistema BA DATA desde el 01-01-2020 a la fecha de redacción de este informe, pueden enumerarse:

- publicación de nuevos datasets
- automatización de datasets relacionados a COVID-19
- publicación de nuevas historias en la sección “historias con datos”.

Durante el año 2021 se actualizó la tecnología del portal, con las características que fueron explicadas en el punto 5.1.2.

Para los años 2021-2022, se planificó agregar nuevas funcionalidades al portal de BA Data. Específicamente, previsualizaciones de los datasets en tablas y gráficos. Otro punto central será organizar las métricas de visitas al portal, descargas de datasets e información relacionada con las actualizaciones, desactualizaciones, frecuencias de cambio, calendarios de publicación, etc. Estos puntos serán publicados para que sean accesibles a los usuarios externos.

También se desarrollarán y publicarán más APIs para que el consumo de datos sea preferentemente por este medio y no por datasets publicados.

¹² Redmine es una herramienta para la gestión de proyectos, que con sus diversas funcionalidades permite a los usuarios de diferentes proyectos realizar el seguimiento y organización de los mismos. Además es posible optimizar su funcionamiento agregando funcionalidades. Incluye un sistema de seguimiento de incidentes con seguimiento de errores. <https://www.redmine.org/> [Accedido el 18/03/2022].

6.4 - Metodología y entornos de desarrollo.

Dentro de la DGARD, la Gerencia Operativa de Ingeniería de Datos (GOID) tiene a su cargo el desarrollo, mantenimiento y soporte de varias aplicaciones y plataformas, entre ellas BA Data. Cada una de estas aplicaciones tiene su proyecto correspondiente en la herramienta Redmine. Todos los requerimientos, tanto de usuarios como internos, se documentan en tickets de Redmine, asociados al proyecto correspondiente, generando un backlog¹³ de peticiones para esa aplicación.

El equipo de desarrollo trabaja con metodologías ágiles, en sprints de scrum¹⁴, ejecutando dos sprints por mes: uno en la primera quincena y otro en la segunda.

Cada sprint se inicia seleccionando del backlog de requerimientos aquellos que formarán parte del sprint¹⁵, según prioridades acordadas entre los usuarios y el equipo de desarrollo. Ello se documenta con el requerimiento funcional cargado en un ticket de la plataforma de gestión de tickets Redmine.

Dicha selección se comunica al equipo de PMs (*Project Manager*)¹⁶ de la Subsecretaría (SSPPBE). Los desarrollos relacionados específicamente con BA Data son comunicados al coordinador del portal BA Data.

Durante las dos semanas que dura el sprint, el equipo de desarrollo se reúne a diario para coordinar las tareas asignadas a cada colaborador y destrabar los temas que requieran una gestión específica.

Al final del sprint se implementan los cambios realizados en cada aplicación, y se comunica al equipo de PMs los requerimientos efectivamente resueltos y aquéllos que por diferentes motivos no hayan podido completarse.

¹³ Backlog es uno de los elementos dentro de la metodología scrum que consiste en una lista de alto nivel con los requerimientos del sistema que se mantiene a lo largo de todo el proyecto, al que se le van modificando los componentes a medida que avanza el proyecto. <https://www.techopedia.com/definicion/13686/scrum> [Accedido el 18/02/2022]

¹⁴ Scrum es un proceso de trabajo para proyectos que permite obtener el mejor resultado posible mediante una estrategia de desarrollo incremental basada en equipos autoorganizados. [https://en.wikipedia.org/wiki/Scrum_\(software_development\)](https://en.wikipedia.org/wiki/Scrum_(software_development)) [Accedido el 03/02/2022]

¹⁵ Un sprint de scrum es un ciclo de trabajo regular y repetible en la metodología de scrum durante el cual el trabajo se completa y se prepara para su revisión. Los sprints Scrum son unidades básicas de desarrollo en la metodología Scrum. Generalmente, los sprints de scrum duran menos de 30 días. <https://www.techopedia.com/definicion/13687/scrum-sprint> [Accedido el 18/03/2022]

¹⁶ Un PM o gerente de proyecto es la persona responsable de liderar un proyecto desde su inicio hasta su ejecución. Esto incluye la planificación, ejecución y gestión de las personas, los recursos y el alcance del proyecto y tiene plena responsabilidad y autoridad para completar el proyecto asignado. <https://www.techopedia.com/definicion/677/project-manager-pm> [Accedido el 18/02/2022]

Finalizado el ciclo de desarrollo, prueba e integración se efectúa la implementación en los sucesivos ambientes: QA ¹⁷ , Homologación ¹⁸ , Producción¹⁹.

Las adecuaciones y nuevos desarrollos se acuerdan con DGCIGA y se canalizan por medio de reuniones y requerimientos en la plataforma Redmine. El equipo planifica los desarrollos en sprints y se aprueban conjuntamente con DGCIGA antes de implementarlos en Producción. La conformidad para la implementación se documenta por correo. Adicionalmente, los tickets de cambio de Redmine se cierran una vez aceptados por el requirente.

6.5 Control de cambios y versionados.

El procedimiento para el control de cambios del desarrollo por parte de la DGARD es el siguiente:

1. Se establecen actualizaciones, mejoras, y/o modificaciones.
2. Se generan los cambios localmente desde un repositorio específico de la ASINF, trabajando siempre en una rama que determine las acciones que se están realizando; si no existe, se crea una nueva.
3. Se inicia el test de prueba.
4. En caso de cumplir con los requerimientos, se pasan los cambios a la rama de producción.

Actualmente el proyecto no posee un control de versionados de software.

¹⁷ El aseguramiento de la calidad (QA) es el proceso de verificar si un producto cumple con las especificaciones requeridas y las expectativas del cliente. El control de calidad es un enfoque basado en procesos que facilita y define objetivos relacionados con el diseño, desarrollo y producción de productos. El objetivo principal de QA es rastrear y resolver las deficiencias antes del lanzamiento del producto. <https://www.techopedia.com/definition/9038/quality-assurance-qa> [Accedido el 18/03/2022]

¹⁸ El entorno de Homologación (o preproducción) sirve para realizar la prueba de estrés final de la "próxima" versión del producto antes de que entre en funcionamiento. Cuando una característica determinada se verifica lo suficiente como para alcanzar la aprobación, se puede mover del entorno de prueba al entorno de Homologación antes de que se lance en el entorno de producción. <https://www.techopedia.com/definition/8989/production-environment> [Accedido el 18/03/2022]

¹⁹ El entorno de producción es un término utilizado para describir el entorno en el que los usuarios finales ponen realmente en funcionamiento el software y otros productos para los usos previstos. <https://www.techopedia.com/definition/8989/production-environment> [Accedido el 18/03/2022]

6.6 - Usuarios del sistema.

Al tratarse de un portal web, los usuarios destinatarios del sistema (ciudadanos, instituciones y empresas), no necesitan tener un usuario para acceder a la información.

Sí están definidos los usuarios para la administración del sistema, los cuales son generados por los miembros del equipo técnico de DGARD. El pedido es realizado formalmente mediante un ticket Redmine. No se otorgan usuarios en el sistema de BA Data a personas que no desarrollen tareas en la Gerencia. Sólo tienen acceso los miembros de la Gerencia Operativa de Ingeniería de Datos (GOID) de la DGAR, accediendo con su nombre de usuario y contraseña.

La Gerencia Operativa de Gobierno Abierto (GOAB) de la DGCIGA, en la sección de “Historias con Datos” también cuenta con un usuario.

6.7 - Roles y Permisos.

En cuanto a los permisos de las reparticiones, la DGCIGA tiene el acceso a Drupal y puede cargar “Historias con Datos”. Tanto la DGCIGA como la DGARD son webmaster de Drupal.

En el site donde están alojadas las otras secciones del portal: “datasets” y “APIs” se utiliza CKAN²⁰ Andino y es administrado por la Dirección General Arquitectura de Datos (DGARD).

La DGARD es la única que posee permisos sobre el CKAN. Existen 2 roles: usuario y administrador.

El rol de usuario solamente tiene acceso a los datasets, archivos, APIs y documentos publicados.

El rol de administrador ingresa mediante una contraseña que posee la DGARD y sus funciones principales son: creación de datasets, publicación de documentación, creación de nuevas organizaciones, acceso al “borrador” donde es posible dejar los datasets que aún no fueron publicados.

No se realizan procedimientos de alta o baja de roles; se utiliza un único rol, tanto de usuario como de administrador.

²⁰ Comprehensive Knowledge Archive Network (CKAN) es una aplicación web de código abierto para el almacenamiento y la distribución de los datos, tales como hojas de cálculo y los contenidos de las bases de datos. Se trata de software libre distribuido bajo la licencia GNU Affero General Public License. <https://ckan.org/> y <https://datos.gob.ar/acerca/ckan> [Accedidos el 18/03/2022]

6.8 - Sesiones, autenticaciones y accesos al Sistema.

La validación de usuarios se realiza a través de la autenticación de Django²¹ para los usuarios con roles de usuario y administrador; y para el proceso de actualización de datos se hace mediante una API REST²².

6.9 - Medios de acceso al sistema.

Para acceder a la información publicada, como medio de acceso al sistema cualquier persona utilizará navegador web a través de internet, de forma completamente pública. No requiere software cliente.

Con relación a la información que debe ser suministrada por las reparticiones, los responsables de tal tarea en éstas deben acceder a una carpeta compartida mediante FTP²³. Así, los archivos se publican mediante FTP. Las áreas de Gobierno suben archivos para actualizar a un FTP desde la red interna del GCABA.

Luego de ello un proceso de actualización automática se conecta desde la LAN con un cliente de desarrollo propio realizado en Python, copia los archivos subidos al FTP a otros servidores, que a su vez lo publican de forma segura en internet.

6.10 Manuales de usuario.

El portal BA Data tiene publicados manuales, instructivos, guías para los/las usuarios/as en la sección de “recursos” <https://www.buenosaires.gob.ar/datosabiertos/recursos>.

Allí se pueden encontrar las siguientes guías:

1- Guía para la publicación de datos abiertos: Conjunto de recomendaciones y buenas prácticas para la publicación de datos en formato abierto.

²¹ Django es un framework de desarrollo web de código abierto, escrito en Python y tiene la finalidad de facilitar la creación de sitios web complejos <https://www.djangoproject.com/> [Accedido el 04/02/2022]

²² API REST es un estilo de arquitectura software para sistemas hipermedia distribuidos como internet. <https://www.ibm.com/ar-es/cloud/learn/rest-apis> y https://es.wikipedia.org/wiki/Transferencia_de_Estado_Representacional [Accedidos el 04/02/2022]

²³ El Protocolo de transferencia de archivos (FTP) es un protocolo cliente/servidor de Internet utilizado para transferir archivos hacia o desde una computadora host. El FTP se puede autenticar con nombres de usuario y contraseñas, está definido en la RFC 959. <https://www.techopedia.com/definicion/1872/file-transfer-protocol-ftp> [Accedido el 18/03/2022]

2- Guía para la identificación y uso de entidades interoperables: Lineamientos y estándares para nombrar entidades en activos de datos, usando sus nombres y códigos oficiales.

3- Guía para el uso y publicación de metadatos: Recomendaciones para las dependencias del Gobierno de la Ciudad a fin de que instrumenten políticas de apertura de datos públicos.

4- Estándar Open Contracting: Documentación de la publicación del dataset de Compras y Contrataciones bajo el estándar de Open Contracting Partnership.

Asimismo, también se cuenta con una guía para principiantes e instructivos para usuarios:

<https://www.buenosaires.gob.ar/datosabiertos/bienvenidoa-al-portal-de-datos-abiertos-de-la-ciudad>

6.11 - Infraestructura, equipamiento, entornos, conexiones.

En sistema se encuentra contenido en diferentes VMs ²⁴ (*Virtual Machines*), que son provistas por la Dirección General de Infraestructura (DGASINF) de la ASINF en modalidad *housing*.

Todos los entornos (Desarrollo, QA, Homologación, Producción) se encuentran separados física y lógicamente por Firewalls (DMZ). Las áreas responsables son la Dirección General de Infraestructura (DGASINF) y la Dirección General de Seguridad Informática (DGSEI), ambas de la ASINF.

6.12 - Tecnologías del Sistema.

Como se indicó previamente el sitio utiliza CKAN Andino y es administrado por la Dirección General Arquitectura de Datos (DGARD).

En el año 2021 el sitio fue migrado a una versión mejorada de CKAN. También, la SSPBE efectuó la actualización a Python, para cumplir con requerimientos de actualización de software por parte de la Agencia de Sistemas de Información (ASI).

La estética del sistema no sufrió grandes modificaciones, y actualmente es la misma de las otras páginas de gobierno y es establecida por la Subsecretaría Experiencia Digital (SSED).

²⁴ Una VM es una aplicación de software que realiza la mayoría de las funciones de una computadora física, y en realidad se comporta como un sistema informático separado. <https://www.techopedia.com/definicion/4805/virtual-machine-vm> [accedido el 18/03/2022]

Los servidores tienen instalado el sistema operativo Red Hat Enterprise Linux en versiones actualizadas.

Los lenguajes utilizados en la programación de los desarrollos son: Python, PL/PGSQL y HTML.

La Base de datos utilizada es PostgreSQL/Postgis versión 10.16.

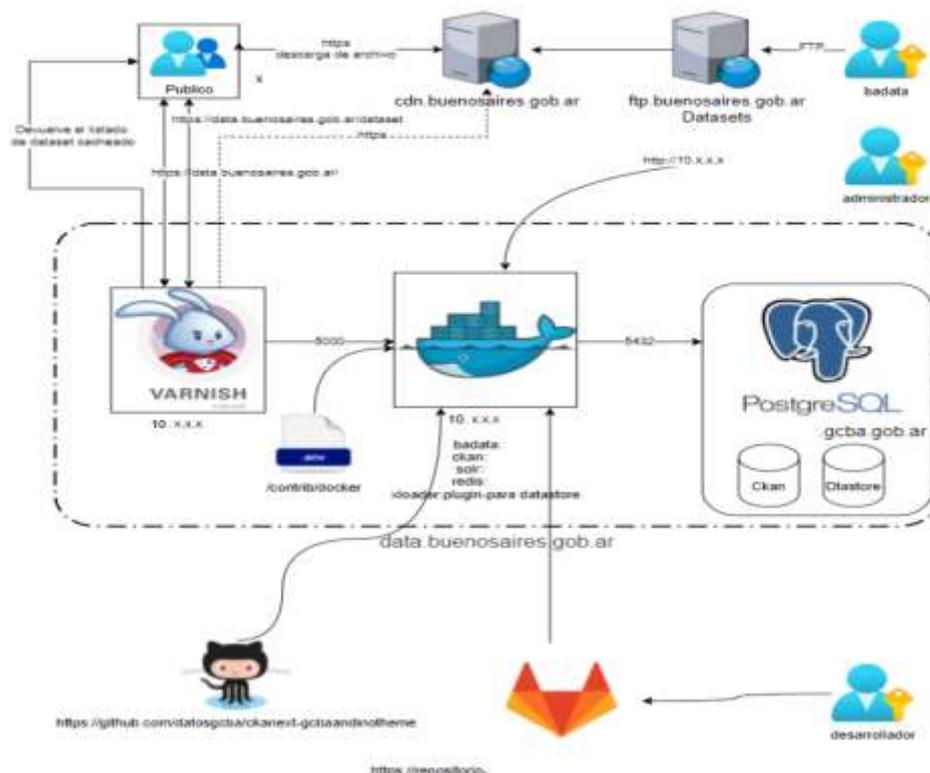
En cuanto a los repositorios, la solución utiliza un contenedor Docker y los files systems provistos por ASIF para el CDN y FTP.

La arquitectura de contenedores incluye estos componentes:

- CKAN: aplicación web.
- DB: base de datos postgres conectada desde un servidor externo.
- Solr: motor de búsqueda
- Redis: motor de base de datos en memoria
- Nginx: middleware

Las versiones y builds de los softwares relevados se encuentran vigentes a la fecha de la redacción del informe.

BADATA Produccion



6.13 Funcionalidades

6.13.1 Funcionalidades. Recopilación de la información.

Los datos que luego serán publicados en el portal BA Data son recibidos por la DGARD a través de los siguientes canales:

a- Por mail: La repartición envía los datos como adjuntos.

Es el método que fue utilizado desde los inicios y actualmente sigue siendo el más empleado.

Cabe señalar que en el momento en que se inició la recopilación de datasets se utilizaba este método. No eran muchos datasets o no requerían actualización seguida, por lo que era más manejable el controlar las actualizaciones, publicar los datos, y solicitar la actualización de los datos cuando estas se vencían. Pero actualmente hay 420 datasets publicados, por lo que cada vez es más difícil que un agente esté continuamente pidiendo esos datos y actualizándolos.

b- En un FTP: La repartición carga el archivo en la carpeta correspondiente del servicio FTP que tiene la DGARD.

Cada día corre un proceso automático, que realiza un control y luego al día siguiente se publican en el portal BA Data. De todos modos, siempre depende que el organismo cargue la información, es decir no es automatizado.

Ejemplo de ello son todos los datos referidos a COVID-19 o calidad del aire. Así desde los Ministerios suben en el FTP los archivos correspondientes y luego del proceso automático, los datos son publicados.

Este es el proceso ideal de envío de información que está tendiendo a implementar como regla la DGARD.

c- Mediante la carga del sistema PGD – Plataforma de Gestión de Datos.

El PDG es un sistema donde pueden cargarse datos de las distintas reparticiones, con diferentes modelos de datos para cada caso.

Operativamente es un sistema web donde se cargan datos en tablas de una base de datos. Tiene distintos módulos para diferentes reparticiones.

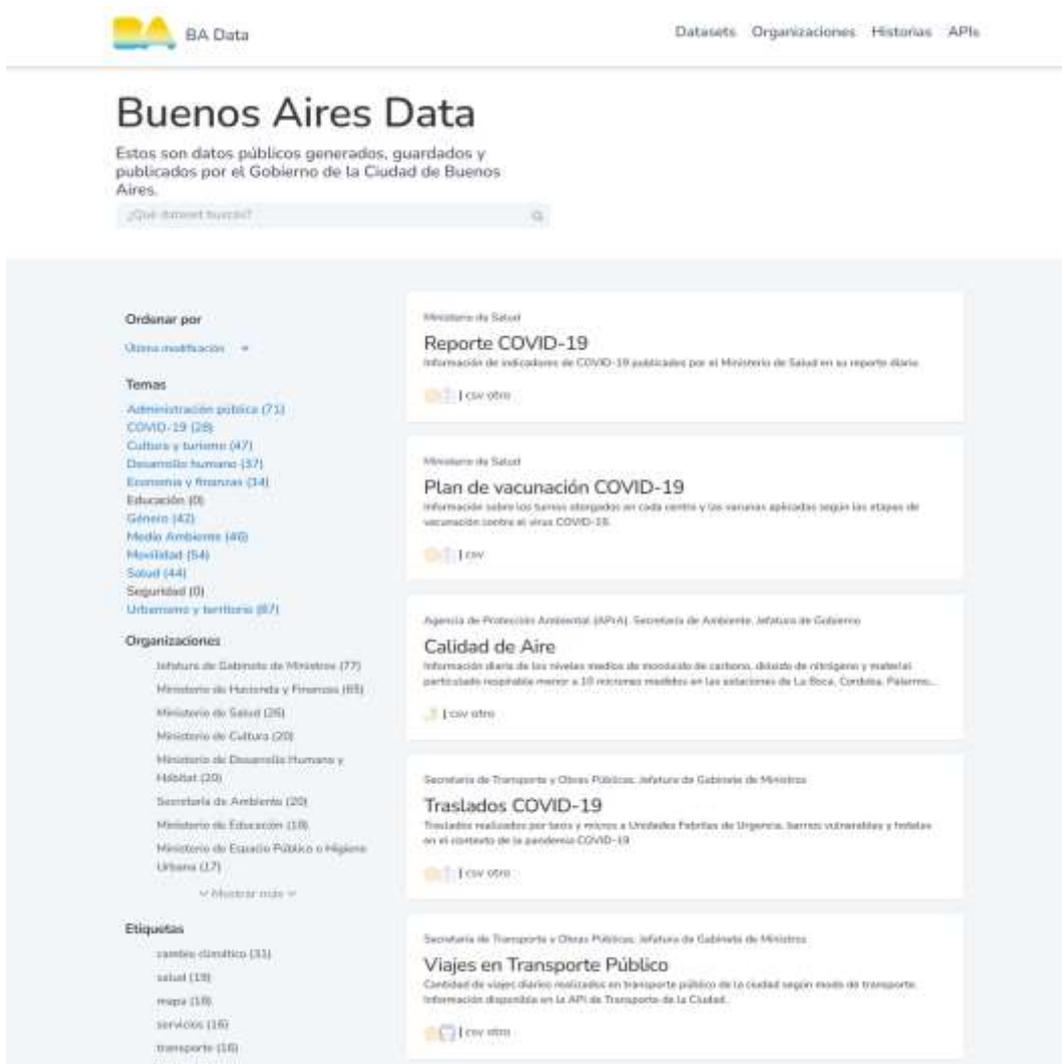
Así por ejemplo, el Ministerio de Cultura envía a la DGARD información específica: el listado de los espacios culturales, con sus datos correspondientes, entre ellos la dirección. Esta información debe ser reflejada en mapas embebidos que están en diferentes páginas del sitio web del GCABA. Para ello la información es cargada al PGD.

La información contenida en el sistema PGD también puede y es descargada y publicada en BA Data.

6.13.2 Funcionalidades. Publicación de la información.

Las funcionalidades que brinda el sistema BA Data son la publicación web de datasets, historias con datos y APIs relacionados con la información generada en diferentes reparticiones del GCABA.

6.13.2.1 Datasets.



The screenshot shows the 'Buenos Aires Data' website interface. At the top, there is a search bar and navigation links for 'Datasets', 'Organizaciones', 'Historias', and 'APIs'. The main heading is 'Buenos Aires Data' with a subtext: 'Estos son datos públicos generados, guardados y publicados por el Gobierno de la Ciudad de Buenos Aires.' Below this is a search input field with the placeholder text '¿Qué dataset buscas?'. On the left side, there are several filter sections: 'Ordenar por' (with a dropdown arrow), 'Temas' (listing categories like 'Administración pública (73)', 'COVID-19 (25)', etc.), 'Organizaciones' (listing government departments like 'Jefatura de Gabinete de Ministros (77)', 'Ministerio de Salud (26)', etc.), and 'Etiquetas' (listing tags like 'cambio climático (33)', 'salud (13)', etc.). The main content area displays a grid of dataset cards. Each card includes the organization name, the dataset title, a brief description, and a 'CSV OTRO' button. The visible datasets are: 'Reporte COVID-19' from the Ministerio de Salud, 'Plan de vacunación COVID-19' from the Ministerio de Salud, 'Calidad de Aire' from the Agencia de Protección Ambiental (APA), 'Traslados COVID-19' from the Secretaría de Transporte y Obras Públicas, and 'Viajes en Transporte Público' from the Secretaría de Transporte y Obras Públicas.

Imagen de la sección de los datasets publicados.

En la sección “Datasets” del sitio, actualmente pueden encontrarse 420 datasets publicados. Allí pueden buscarse por su nombre, tema, organización que genera la información, o etiquetas (tags) de información.

Las actualizaciones de los datos de los datasets se realizan con

determinada periodicidad. De esta manera, las actualizaciones pueden ser: eventuales, diarias, semanales, quincenales, mensuales, trimestrales, semestrales y anuales.

Cada dataset a su vez cuenta con diferentes “recursos”. Los recursos son diferentes conjuntos de datos del mismo dataset, ya se encuentre la información en diferentes formatos (.csv, .xlsx, Json, GeoJson, shape, etc) o referida a cortes de información específicos: ej. referida a un año en particular (2019, 2018, etc).

La información contenida en los datasets es provista por las diferentes reparticiones del GCABA, es la DGARD la encargada de recibirla, controlarla y publicarla.

Se realizó un relevamiento de los datasets que se encuentra en el punto 8 de este informe.

6.13.2.2 Sección “Historias”.

“Año del 40º Aniversario de la Guerra de Malvinas. En homenaje a los veteranos y caídos en la defensa de las Islas Malvinas y el Atlántico Sur”

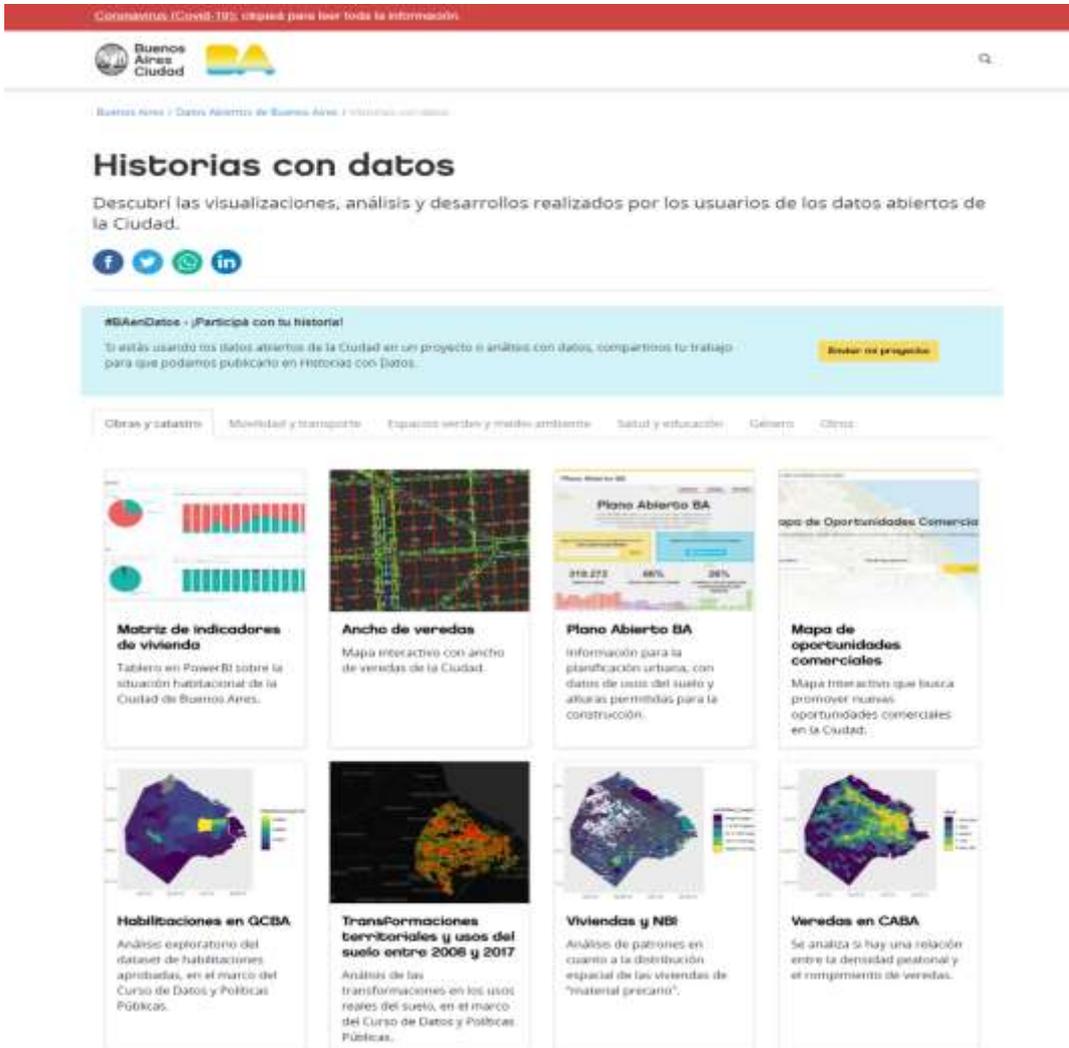


Imagen de la sección Historia con Datos, donde pueden verse los links a desarrollos realizados por externos con datos obtenidos en el portal BA Data.

La sección “Historias con Datos” del portal fue desarrollada por iniciativa de la DGCIGA, y es mantenida por ella.

Allí los usuarios del público realizan sus propias apps, las publican y envían el link a la DGCIGA. Ésta analiza la viabilidad y funcionalidad de la app, luego de aceptar el link, y es subido a la sección de “Historias con Datos”.

Esta sección recibió mejoras durante el año 2020. Así toda la información ha sido reubicada y es gestionada mediante la aplicación Drupal. Con anterioridad se encontraba en la aplicación CKAM, que es un desarrollo recibido de Nación. Esto se realizó con el apoyo de un proveedor (Devartis S.A.), teniendo en cuenta que se debía reorganizar cada uno de los contenidos y testear

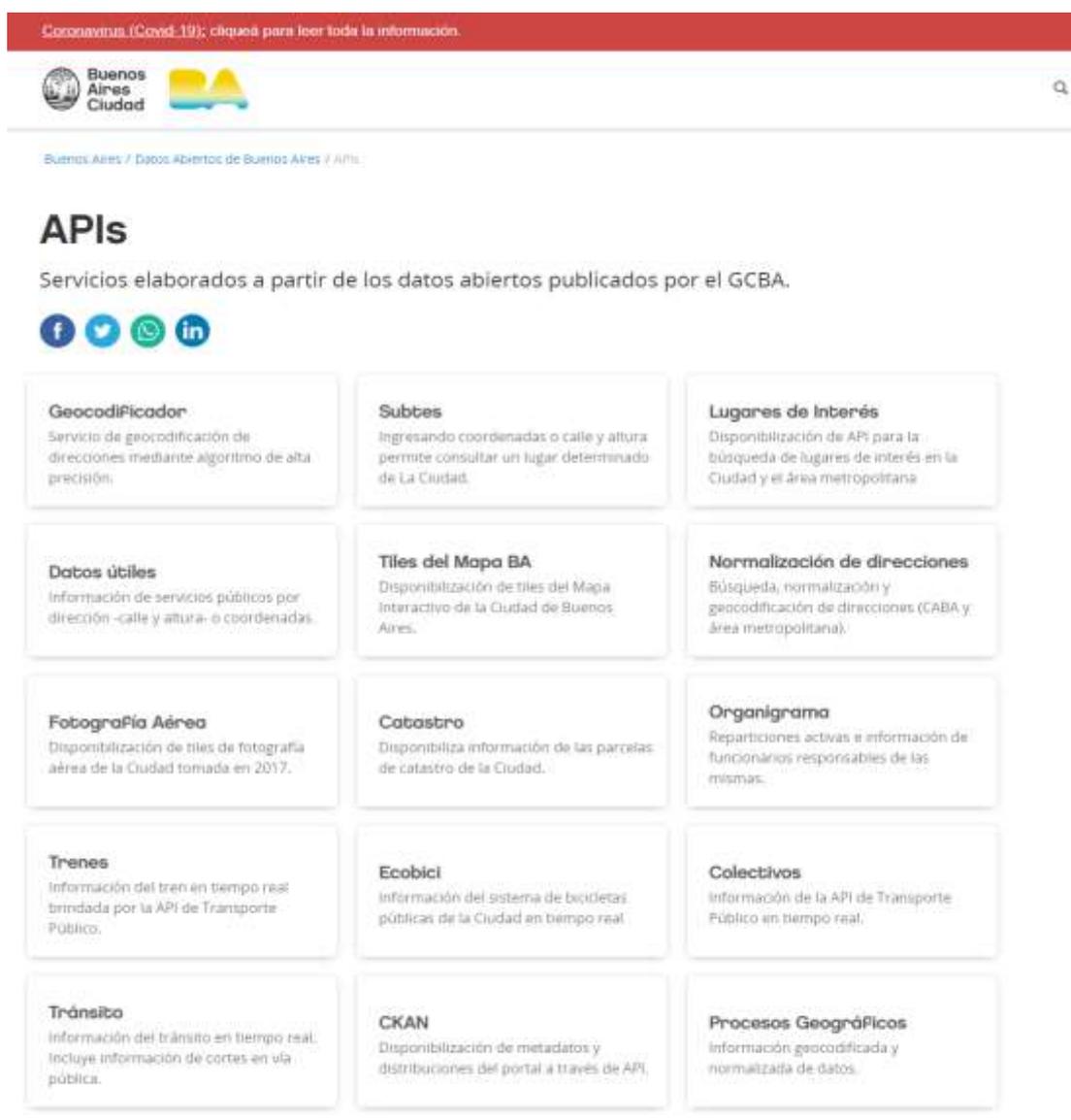
“Año del 40° Aniversario de la Guerra de Malvinas. En homenaje a los veteranos y caídos en la defensa de las Islas Malvinas y el Atlántico Sur”

imágenes, entre otras tareas. La definición de esta reorganización se llevó adelante de manera conjunta con la DGARD.

La tarea realizada fue pasar a Drupal toda la sección de historias con datos. Se crearon más de 100 páginas internas y tarjetas, así como configurar las imágenes con tamaño correcto correspondiente.

La DGARD fue quien definió utilizar Drupal como mejor práctica para gestionar (editar) el contenido.

6.13.2.1 APIs.



Coronavirus (Covid-19): clickeá para leer toda la información.

Buenos Aires Ciudad

Buenos Aires / Datos Abiertos de Buenos Aires / APIs

APIs

Servicios elaborados a partir de los datos abiertos publicados por el GCBA.

f t w in

Geocodificador Servicio de geocodificación de direcciones mediante algoritmo de alta precisión.	Subtes Ingresando coordenadas o calle y altura permite consultar un lugar determinado de La Ciudad.	Lugares de Interés Disponibilización de APIs para la búsqueda de lugares de interés en la Ciudad y el área metropolitana.
Datos útiles Información de servicios públicos por dirección -calle y altura- o coordenadas.	Tiles del Mapa BA Disponibilización de tiles del Mapa Interactivo de la Ciudad de Buenos Aires.	Normalización de direcciones Búsqueda, normalización y geocodificación de direcciones (CABA y área metropolitana).
Fotografía Aérea Disponibilización de tiles de fotografía aérea de la Ciudad tomada en 2017.	Catastro Disponibiliza información de las parcelas de catastro de la Ciudad.	Organigrama Reparticiones activas e información de funcionarios responsables de las mismas.
Trenes Información del tren en tiempo real brindada por la API de Transporte Público.	Ecobici Información del sistema de bicicletas públicas de la Ciudad en tiempo real.	Colectivos Información de la API de Transporte Público en tiempo real.
Tránsito Información del tránsito en tiempo real. Incluye información de cortes en vía pública.	CKAN Disponibilización de metadatos y distribuciones del portal a través de API.	Procesos Geográficos Información geocodificada y normalizada de datos.

Imagen de la sección APIs del portal BA Data donde pueden apreciarse las APIs disponibles.

En la sección “APIs” se encuentran los links de acceso a las APIs publicadas que ofrecen información en tiempo real de áreas del GCABA.

Las API están orientadas principalmente para desarrolladores que desean escribir código que interactúe con los sitios del portal y sus datos.

La mayoría de las APIs publicadas son de Transporte o de Catastro. La idea es seguir sumando APIs, ya sea desarrolladas por la DGARD o distintas áreas del Gobierno.

La API de Transporte fue desarrollada en el Ministerio de Transporte, y está alojada y funciona en su sitio web.

La API de Catastro fue desarrollada y es mantenida por la Dirección General de Ciencias de la Información (DGCINFO), también dependiente de la de la SSPPBE.

En la versión anterior del portal, cuando se quería acceder a las APIs, la página direccionaba a otras APIs vinculadas. Es decir, se encontraba enlazado a otras páginas donde estaban alojadas las APIs.

A partir del mayo/junio 2021 ello fue modificado para que directamente se pueda hacer la consulta a la API desde el portal de BADATA. Las APIs de catastro o datos geográficos que son de la misma Subsecretaría ya funcionan de esta manera y así pueden ser consultadas.

Para el caso de la API de Transporte, para poder hacer la consulta, previamente la API te solicita un ID de cliente y una password. El Ministerio de Transporte requiere la registración previa del usuario mediante mail para utilizar su API. Lo que se está tratando de hacer ahora es que sea una API libre o abierta, como son las de catastro.

Así mismo, como se indicó previamente, el objetivo buscado es seguir sumando nuevas APIs. Aún existen algunas APIs de Catastro que no están publicadas en BA Data; y otras de otras reparticiones de Gobierno. Ya sea porque la cree la DGARD o porque se convenga con otras reparticiones que publican datos en BA Data y puedan generar sus propias APIs.

El equipo de Auditoría realizó un relevamiento de la operatividad de algunas las APIs publicadas utilizando el software Postman. La conclusión fue que las APIs publicadas se encuentran funcionando, los enlaces operativos y la información devuelta por la API era la esperada.

6.13.3 - Funcionalidades. Usabilidad y experiencia del usuario (UX).

El uso del sitio es sencillo e intuitivo y el aspecto estéticamente agradable. Intenta otorgar una buena experiencia al usuario del público que accede a la web. Se utiliza el recurso estético de *cards* en la sección de “Historias con Datos”.

Es fácilmente entendible qué información tiene cada sección, cómo accederla, y las características de cada dataset, recurso, “Historia”, y cómo interactuar con las APIs publicadas. Cuenta con la cantidad adecuada de links, y se entiende que realiza cada uno.

No obstante lo anterior, durante el relevamiento del portal web, el equipo de auditoría pudo comprobar la existencia de ciertos errores en las funcionalidades:

- Al navegar dentro de un dataset seleccionado, al acceder al link de “ver esta versión”, el sistema da un error y no muestra información.
- Al acceder al link de “cambios”, la nueva página trae la información de los cambios pero no funciona el botón “Show metadata diff”.
- En la lista de “Organizaciones”, al hacer click en una de ellas, el número de datasets que se listan no coincide con el número anteriormente informado entre paréntesis Ej. la organización “Jefatura de Gabinete de Ministros” informa que tiene 113 datasets, pero al accederla, solo se encuentran listados 12.

6.14 Controles.

No existe un control sistemático de todos los datasets que se reciben y son publicados. Los que son enviados por las áreas, pasan por un control sobre la codificación de los datos (tienen que venir en UTF-8), la cantidad de campos y sus nombres. En ocasiones se detectan observaciones sobre los tipos o formatos de datos. Actualmente la DGARD está trabajando para que esta validación sea automática.

Los datos que sí pasan por una validación previa y control sistematizado son los que se publican a partir de un proceso de integración ETL²⁵; por ejemplo los datos COVID-19. Este proceso controla el tipo de datos y cantidad de campos.

²⁵ ETL: Extract, Transform and Load (extraer, transformar y cargar). Es el nombre que se le da a un proceso que extraiga datos de una fuente, luego haga ciertas operaciones sobre los datos y por último los deposite en otro repositorio, BBDD, etc.
https://www.sas.com/en_us/insights/data-management/what-is-etl.html [Accedido el 03/02/2022]

Tampoco existe un control automático sobre la actualización o no de los archivos provistos por las reparticiones, ni sobre la actualidad de la información publicada.

6.15 - Reportes, Notificaciones y Avisos.

Desde el sistema solamente se puede generar el Catálogo de Datasets, el cual informa la cantidad de datasets, sus responsables, frecuencia de actualización y fecha de últimas modificaciones.

Además, del Catálogo de Datos, se utilizan servicios externos para medir el desempeño: Google Analytics, que mide las visitas al portal de datos y los datasets; y el Kibana, que mide las descargas de archivos del portal.

El sistema por ahora no cuenta con notificaciones o avisos automáticos. Está previsto generar notificaciones cuando se logre una integración con OpenShift²⁶, ya que este último sistema cuenta con notificaciones y avisos que pueden implementarse para el portal.

6.16 - Integración del Sistema.

No hay integración con otros sistemas del GCABA. Los datos que se publican son producidos por las distintas áreas de Gobierno y el sistema se dedica a publicar los datos que producen estas áreas.

Si bien no se tratan de integraciones, existen vinculaciones con APIs desarrolladas por algunas pocas reparticiones del GCABA, como se ha señalado anteriormente.

6.17 - Soporte Técnico.

Entre las funciones de la DGCIGA y DGARD se encuentra la de realizar el soporte a los usuarios externos sobre datasets, recursos, APIs y guías.

Además, los usuarios (ciudadanos) pueden pedir apertura de datos mediante pedidos de acceso a la información pública (Ley CABA 104).

Se cuenta con un mail donde se reciben las consultas: gobiernoabierto@buenosaires.gob.ar, administrado por la DGCIGA. Además, la Dirección General de Arquitectura de Datos (DAGARD) cuenta con un mail al cual llegan consultas: datosabiertos@buenoaires.gob.ar.

²⁶ <https://www.redhat.com/en/technologies/cloud-computing/openshift> [Accedido el 03/02/2022]

Por otra parte, también ingresan consultas a través de DM (direct message) de Twitter y otras redes sociales a la DGCIGA. Las consultas y pedidos que se reciben por diversos medios son tratados en todos los casos.

Los casos de soporte referidos a asistencia y a ayuda sobre la información publicada son tratados por la DGCIGA. Los casos que requieran soluciones técnicas, o sean sobre errores en los datos de los datasets, o problemas con su calidad o si no funcionan; son derivados y/o resueltos por la DGARD.

En caso de que lo solicitado sea la apertura de un nuevo dataset, la DGCIGA y la DGARD trabajan en conjunto con el objetivo de poder darle apertura a esa información.

El procedimiento utilizado es que la DGCIGA crea tickets o reenvía el mail; y la DGARD sigue el caso. Así quedan documentados los pedidos de los usuarios.

Las áreas internas del GCABA también pueden generar tickets con respecto a la información que provean. Así pueden solicitar aperturas, actualizaciones y controles en los datos que se publican.

Los tickets se generan en la plataforma Redmine. El sistema de tickets está reservado a las áreas internas de Gobierno. No los pueden crear los usuarios del público.

6.18 - Gestión de la Información. Protección de datos personales.

La información publicada en BA Data es pública y provista por las áreas que la generan o la compilan. El equipo de la Gerencia Operativa de Integridad de Datos (GOID) de la DGARD realiza una revisión de los datos al momento de publicarlos para asegurar que no se publiquen datos con otros niveles de privacidad, y que mantengan integridad respecto de los datos que se vienen publicando con anterioridad.

En cuanto a la definición de la propiedad y guarda de la información que será publicada, el área propietaria del dato es la repartición que provee la información y el área custodia es la DGARD, y se definen al acordar la publicación de cada *dataset*.

En ocasiones sucede que las áreas cambian de nombre o de lugar en el organigrama del GCABA, o bien hay rotación de personal en un área y cambia quien emite o proporciona el *dataset*, quedando esta información desactualizada. Durante 2021 se llevó a cabo un relevamiento y actualización de los referentes, tanto de las áreas productoras de los datos como de aquellas que los reúnen y los proporcionan a BA Data.

Con respecto a la protección de los datos personales, vale aclarar que no se publican datos personales en los *datasets*. Los únicos dos que cuentan con

datos personales son los de “Declaraciones Juradas” y “Sueldos de funcionarios”. Ambos solamente publican información de funcionarios de Gobierno de nivel ministerial -hasta Secretario/a y no más abajo-. Los datos personales publicados en dichos *datasets* son nombre y apellido, CUIL e información relacionada a sus salarios, asignaciones y sus bienes. Dicha publicación se realiza en cumplimiento a las normativas específicas sobre ética pública.

Es el área que comparte los datos es quien generalmente define los plazos de actualización de la información. La DGARD realiza sugerencias sobre la periodicidad. Existe un tipo de frecuencia que se llama “eventual”, definida para los datos que no reciben cambios por largos períodos de tiempo. Por ejemplo: iglesias en la Ciudad o arroyos, no requieren tener una frecuencia.

Los criterios y quienes son los funcionarios responsables de la selección de qué información recopilada será objeto de publicación depende del proceso de apertura: En un caso son las áreas las que proponen los datos que desean abrir en el portal. En otro, son la DGARD y/o DGCIGA quienes solicitan la apertura a ciertas reparticiones dueñas de datos que se estiman estratégicos.

Cualquiera de las dos modalidades implica un control en el área de los técnicos especialistas en datos y, si se trata de información sensible, de los responsables legales de la SSPPBE. La repartición que abre la información es la que debe garantizar que los datos sean actualizados en la frecuencia establecida

6.19 Satisfacción del usuario.

La Dirección General de Calidad Institucional de Gobierno Abierto (DGCIGA) no cuenta -por el momento- con sistema o medio estandarizado que permita conocer la satisfacción del usuario. No obstante, la DGCIGA analiza la calidad de los productos y satisfacción de los usuarios a través de las redes sociales, en especial con una gran interacción de comunicación y participación con la comunidad de usuarios de datos, e incluso mediante comentarios y pedidos.

Puede señalarse que, la DGCIGA, utilizando twitter y otras redes sociales, realiza desafíos de visualizaciones de datos generados por el GCABA. Así, por ejemplo, se establece la consigna de utilizar un dataset en particular y generar una visualización o historia específica acorde. Junto con DGARD, de manera conjunta, definen cuales son las propuestas recibidas más importantes para publicarlas en historias.

6.20 Estándares.

En cuanto a los estándares, dentro de la solapa “guías”, se encuentran todas las especificaciones que las reparticiones que hagan datasets en el Gobierno deberían seguir. En caso de que no se sigan estos estándares, la Dirección General de Arquitectura de Datos (DGARD) es la encargada de estandarizar y que cumpla con dichos estándares con el objetivo de que los datos sean interoperables y se puedan integrar.

Así mismo, la DGCIGA mantiene vínculos con algunos organismos que han definido estándares internacionales que han sido adoptados por reparticiones del GCABA.

Es el caso del dataset de compras abiertas del BAC (buenos aires compras). El mismo sigue un estándar internacional que se llama OCDS que lo establece la organización internacional Open Contracting Partnership. Este organismo realiza sobre la DGCIGA un seguimiento respecto del cumplimiento de ese estándar. La DGCIGA promovió la implementación de este estándar y la DGARD es la encargada de actualizar ese dataset en conformidad.

7.- Seguridad sobre el Sistema y Contingencia.

7.1.- Política de Resguardo (Backup) y recuperación (Restore).

La Agencia de Sistemas de Información (ASINF) es quien realiza la política y acciones de backup y restore.

Cómo se indicó previamente, el sistema de BA Data está contenido en VMs que son hospedadas por la ASINF.

Conforme al procedimiento para respaldo de servidores virtuales de la ASINF, la totalidad de las VMs en Vcenter son añadidas de forma automática a rutina de backup debido a las políticas predefinidas en las tareas de respaldo diarias. El software de respaldo utilizado en éste caso es “Veeam Backup & replication”²⁷.

Los Backup perduran durante el tiempo definido en las políticas de retención de ASI y luego se expiran mediante procesos internos de depuración. Si una VM es eliminada, la tarea de backup la identifica y excluye automáticamente del backup.

La información que se resguarda es el servidor virtual en su totalidad.

²⁷ Veeam Backup & Replication es una aplicación de respaldo de información utilizada para entornos virtuales, la cual permite verificar automáticamente la recuperación de la información almacenada. <https://www.veeam.com/>.

El tiempo de guarda para el software Veeam son los siguientes:

- Backup diarios, se ejecutan una vez al día y tienen una retención de 30 puntos de restauración.
- Backup mensuales, consisten en la copia a cinta del backup full correspondiente a los datos del último día del mes. Tienen una retención de 2 años (24 meses).
- Backup anuales, consisten en la copia a cinta del backup full correspondiente a los datos del último día del año. Tienen una retención de 10 años (120 meses).

La ubicación del repositorio de guarda del software Veeam: Los backup diarios son almacenados en repositorios a discos (HDD²⁸ y SSD²⁹) que se encuentran asociados a sus respectivos servidores. Los backup mensuales y anuales son almacenados en cajas ignífugas.

En cuanto a las Pruebas de Restauración, la ASINF informa que dada la regularidad de restauraciones para el entorno virtual (VMs), actualmente se están tomando como pruebas de restauración las mismas solicitudes realizadas por los solicitantes a través del sistema de tickets NOC³⁰.

7.2.- Antivirus, Seguridad Física, Alta Disponibilidad, Plan de Contingencia, Plan de Recuperación.

Los servicios de Antivirus, Seguridad Física, Alta Disponibilidad, Plan de Contingencia, Plan de Recuperación sobre los servidores del sistema es provista por la DGASINF de la ASINF³¹.

La Agencia de Sistemas de Información (ASINF) cuenta con un Data Center sito en Independencia 635 y otro Centro de Cómputos de contingencia, en Uspallata 3160, para poder continuar operando para el caso de que el Data Center de Independencia fuera comprometido. Ambos Data Centers se encuentran protegidos por controles de acceso biométricos y vigilancia; y

²⁸ HDD (Hard Drive Disk, en español Disco Duro) es un tipo de disco integrado por componentes mecánicos que sirve para almacenar los datos en forma permanente, los cuales no se borran cuando la unidad carece de alimentación por energía eléctrica.

²⁹ SSD (Solid State Drive, en español Unidad de Estado Sólido), es un disco que almacena los datos en microchips con memorias flash interconectadas entre sí.

³⁰ Un sistema de tickets NOC (Network Operation Center o centro de operaciones de red) es una herramienta esencial de la gestión de TI moderna al centralizar las operaciones clave en una sola ubicación. <https://www.ninjaone.com/blog/network-operations-center-noc-explained/> [accedido el 18/03/2022].

³¹ La AGCBA realizó el informe “Agencia de Sistemas de Información” cuyo N° de Proyecto es 10.21.03, sobre las conexiones y operaciones de la ASINF, donde podrá verse en mayor profundidad los servicios de operaciones brindados por la DGIASINF.

cuentan con sistemas contra incendio y controles ante fallos de energía (UPS y generadores).

Plan de contingencia: La DGASINF de la ASINF no contempla un plan desarrollado específicamente para las VMs del portal de BA Data. La solución de contingencia por *default* es proveer los respaldos (backup) en caso de que se solicitaran los mismos.

7.3.- Registro de Eventos (logs³²).

El portal cuenta con logs de trazabilidad que permiten obtener información de accesos, actividades y transacciones como las descargas de datasets. Dichos logs actualmente se recopilan con la herramienta ELK³³ y la información de descargas de datasets se visualiza en un tablero de Kibana³⁴ con refresco cada 3 semanas. Actualmente se está en pleno proceso de centralización de los logs en un Data Warehouse³⁵, para obtener métricas de períodos mayores de tiempo.

Asimismo, el portal cuenta con un log de actualización de datasets. Esto lo genera y almacena un proceso desarrollado en python³⁶ que es el que realiza las actualizaciones de los datasets.

8. Relevamientos de los datasets publicados.

El equipo de auditoría realizó dos relevamientos de la información contenida en los datasets publicados en la solapa “datasets” del portal de Buenos Aires Data.

En el primero se relevó la totalidad del universo: 423 datasets. Se agruparon en 3 conjuntos: “Eventual”, si el dataset no tiene una periodicidad para su actualización; “Actualizado”, si la fecha de actualización estaba dentro del plazo para hacerlo; y “Desactualizado”, si no lo estaba. Con ello se confeccionaron los siguientes cuadros y gráfico:

³² El registro de datos es el proceso de recopilación y almacenamiento de datos durante un período de tiempo para analizar tendencias específicas o registrar eventos/acciones basados en datos de un sistema, red o entorno de TI. <https://www.techopedia.com/definition/596/data-logging> [accedido el 18/03/2022]

³³ <https://www.elastic.co/es/what-is/elk-stack> [Accedido el 03/02/2022]

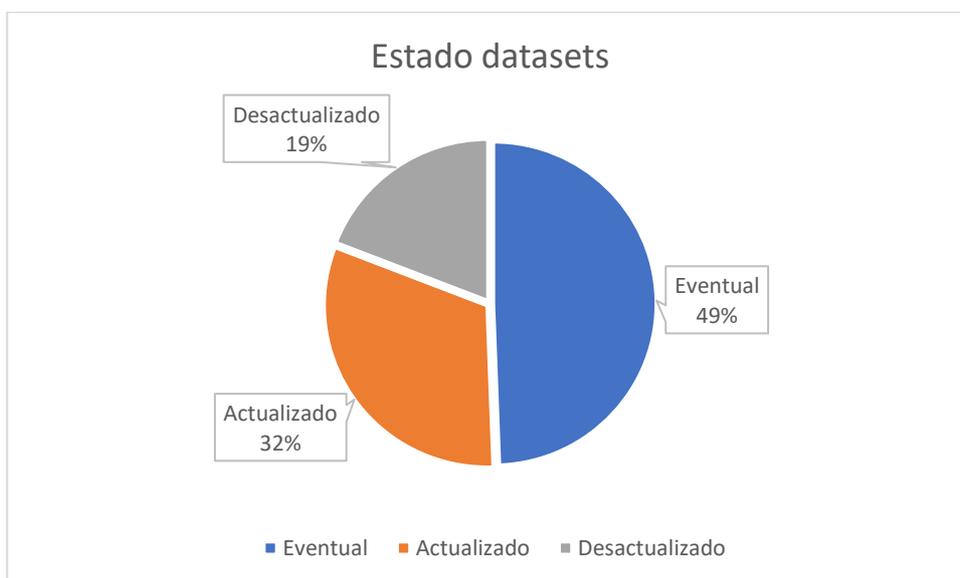
³⁴ <https://www.elastic.co/es/kibana/> [Accedido el 03/02/2022]

³⁵ <https://platzi.com/blog/ques-un-data-warehouse/> [Accedido el 03/02/2022]

³⁶ <https://www.python.org/> [Accedido el 03/02/2022]

“Año del 40° Aniversario de la Guerra de Malvinas. En homenaje a los veteranos y caídos en la defensa de las Islas Malvinas y el Atlántico Sur”

Estado	Cantidad Dataset	%
Eventual	209	49%
Actualizado	133	31%
Desactualizado	81	19%
Total	423	100%



De lo anterior se desprende que:

- El 49% de los datasets son de actualización eventual, se trata de información que no es dinámica por lo que se ha definido que no necesita ser actualizada por plazos determinados. Ej. datasets “Hospitales”, donde están listados los hospitales del GCABA.
- El 51% de los datasets tienen plazo de actualización, y dentro de ellos un 31% se encuentra actualizado; y un 29% desactualizado, o sea no se cargaron datos luego de vencido el plazo para su actualización.

Desagregando los datasets definidos como de actualización periódica, encontramos en cuanto a su frecuencia de actualización los siguientes valores:

“Año del 40° Aniversario de la Guerra de Malvinas. En homenaje a los veteranos y caídos en la defensa de las Islas Malvinas y el Atlántico Sur”

Frecuencia Actualización	Cantidad Datasets	%
cada 2 años	2	0,93%
anualmente	84	39,25%
semestralmente	33	15,42%
cada 3 meses	34	15,89%
cada 2 meses	1	0,47%
mensual	37	17,29%
quincenal	3	1,40%
semanal	2	0,93%
diaria	11	5,14%
continua (API)	7	3,27%
Total	214	100,00%

En un segundo relevamiento se procedió a un análisis de la información contenida en los archivos de los datasets, sobre un muestreo mínimo de 5 casos que cumplieran con los siguientes criterios: i- que estuvieran dentro del grupo de los datasets actualizados del cuadro anterior, ii- que contuvieran una columna referida a la fecha del dato, iii- que fueran de frecuencia de actualización menor o igual a la semestral, iv se prefirió los recursos de tipo .csv sobre otros (ej. xlsx). El resultado es el siguiente cuadro.

Relevamiento de datos contenidos en datasets de actualización periódica (efectuado el 17/03/2022).

#	Dataset	Frecuencia actualización	Recurso utilizado	Fecha de último registro	Fecha actualización dataset	Estado de la Información
1	Permisos Eventos Masivos	Mensual	permisos-eventos-masivos.csv	27/10/2019	4/2/2022	desactualizada y solo incluye 6 meses del año 2019
2	Flujo vehicular	Diaria	dataset_flujo_vehicular.csv	12/3/2022	16/3/2022	actualizada
3	Llamados 107 - COVID	Diaria	llamados_107_covid.csv	13/12/2021	16/3/2022	poco actualizada
4	Calidad de aire	Diaria	calidad-aire.csv	1/3/2022	16/3/2022	actualizada
5	Dictámenes jurídicos de la Procuración General	Semestral	dictámenes-procuracion-general.csv	27/8/2019	21/9/2021	desactualizada

Del análisis del cuadro puede observarse que si bien en el sitio indica una fecha de actualización del dataset que se encuentra dentro del plazo esperado para la actualización, un examen del contenido de la información contenida en el recurso del dataset muestra que se encuentra actualizada (2 casos), poco actualizada (1 caso) o desactualizada (2 casos). Incluso la información puede abarcar solo un periodo de tiempo y nunca haber sido actualizada, como parece indicarlo el dataset “Permisos Eventos Masivos”.

Seguramente ello sea consecuencia de que la rutina de actualización automática desarrollada por la DGARD se ejecuta correctamente en los plazos correspondientes, pero las reparticiones que deben subir la información no lo efectúan en el plazo correspondiente. Es decir, la rutina se ejecuta correctamente pero no hay nuevos archivos que actualizar.

Se concluye asimismo que los datasets relacionados con los temas de transporte, cambio climático y covid-19 cuentan con la información más actualizada.

Integridad y calidad de la información: En esta Auditoría no se realizó un estudio general de la integridad y calidad de la información contenida en los datasets. No obstante, al ir revisando los recursos de los datasets, se encontraron casos donde:

- En campos de texto, en las palabras con tildes los caracteres estaban reemplazados por otros. Por ejemplo, no fueron generados utilizando UTF-8). Un caso de estas características es el recurso de dataset “dictámenes-procuracion-general.csv”.
- Faltan datos en ciertos campos se encuentran vacíos para algunos registros sin que pueda entenderse que hay una razón para ello. Un ejemplo de ello es el recurso “calidad-aire.csv”, en ciertas columnas hay registros que indican “s/d” (se entiende “sin datos”), pero en otras directamente se encuentran vacías.

VI. OBSERVACIONES

Análisis del grado de cumplimiento del Marco de Referencia COBIT 2019

➤ Alinear, Planificar y Organizar - APO02

- Gestionar la estrategia

1. La información que generan las reparticiones del GCABA se halla distribuida en cada una de ellas, no se encuentra centralizada, vinculada o compartida salvo para casos y sistemas específicos. No existe una política general para integrar la información generada por las reparticiones del GCABA. Así, para publicar en el portal de datos públicos abiertos Buenos Aires Data la DGARD y DGCIGA necesitan caso por caso solicitar a cada repartición que comparta los datos que genera. La definición de una estrategia de gestión integral de la información pública para el GCABA es de suma importancia no solo para el objetivo de un gobierno abierto sino para la toma de decisiones de políticas públicas con base en la evidencia.

➤ Alinear, Planificar y Organizar - APO03

- Gestionar la arquitectura empresarial

2. En principio, cabe señalar que la definición es compleja en cuanto a qué criterios estratégicos deben tomarse para la publicación información generada por las reparticiones del GCABA; y que las áreas intervinientes DGCIGA y DGARD buscan activamente la apertura de datos en poder de otras reparticiones que son demandados por actores de la sociedad civil, academia y prensa.

No obstante, no se encuentra definido el universo de la información pública que generan las distintas reparticiones del GCABA y que luego serían publicadas en el portal de datos abiertos Buenos Aires Data. En su defecto, tampoco existen criterios o definiciones estratégicos de qué información es más relevante para ser publicada en el portal. Definir el universo o criterios de importancia sobre la información a publicar en el portal permite confirmar que la información es la más relevante.

3. La información que va a ser publicada debe ser solicitada para cada caso. El flujo de la información no es continuo ni

automático y ésta se ve afectada por desactualizaciones. Las áreas responsables de la publicación de la información (DGARD – DGCIGA) se ven afectadas por el accionar de las áreas productoras de la información cuando incumplen con la actualización de los datos. Solo pueden indicarle al área que debe cumplir con la frecuencia de actualización, y no pueden garantizar que la información se encuentre actualizada. Asimismo, mucha de la información es de actualización eventual. Contar con información actualizada es uno de los pilares en los que debe sostenerse un portal de gobierno abierto.

4. Existen pocas soluciones (APIs) que permitan la disponibilidad inmediata de la información a medida que es generada. Una mayor disponibilidad de la información redundaría en una mayor calidad de las prestaciones del sistema.

➤ Alinear, Planificar y Organizar - APO07

- Gestionar los recursos humanos

5. Se relevó que la gran mayoría de los agentes que prestan tareas técnicas dentro de la DGARD se encuentran contratados bajo la modalidad de Locación de Obra y Servicios, no perteneciendo a la dotación de agentes del GCABA. La falta de formalización en la relación laboral de los agentes podría generar inestabilidad en la funcionalidad del sistema por la rotación del recurso humano en las tareas asignadas.

➤ Alinear, Planificar y Organizar - APO09

- Gestionar los acuerdos de servicio

6. No posee Acuerdos de Niveles de Servicio (SLA)³⁷ por los servicios tecnológicos recibidos. En especial no cuenta con acuerdos formales sobre condiciones y periodicidad de publicación de la información entre las reparticiones proveedoras de los datos y las encargadas de publicarlos. La carencia de acuerdos formalizados genera una asignación difusa de la responsabilidad en el nivel de los servicios acordados.

³⁷ SLA Service Level Agreement, (en español, Acuerdo de Nivel de Servicio), es un acuerdo firmado entre el proveedor de un servicio y el cliente. El objetivo es establecer el nivel del servicio y sus parámetros, tales como los plazos de entrega, responsables, tiempos de respuesta, tipo de soporte técnico, etc

➤ Alinear, Planificar y Organizar - APO14

• Gestionar los datos

7. Se halló que gran cantidad de los datasets tenían el plazo de la frecuencia de actualización vencido. Asimismo, que los datos contenidos dentro de algunos de los datasets se encontraban desactualizados, con errores y omisiones en la información contenida.
8. No existen controles automáticos por medio de softwares que intenten garantizar que se cumplan los plazos de actualización por parte de las reparticiones, y que la información se encuentre efectivamente actualizada, se encuentre normalizada y no contenga errores o esté incompleta.
9. No se realizan, junto con la ASINF, pruebas de restauración específicas de las VMs que componen el sistema BA Data. Efectuar pruebas periódicas de restauración es un buen indicador de operatividad de las políticas de respaldo de la información.

➤ Construir, Adquirir e Implementar - BAI09

• Gestionar los activos

10. No expone procesos de gestión de activos de información, tales como el propietario del dato, usuario y custodio de la información utilizada. Tampoco procedimientos de evaluación y clasificación de la información según su criticidad, nivel de integridad, confidencialidad, disponibilidad y privacidad.

➤ Construir, Adquirir e Implementar - BAI11

• Gestionar los proyectos

11. No expone los planes de corto, mediano y largo plazo asignados al sistema. Es necesario establecer una estrategia de ejecución de proyectos afectados a la modificación y mejoramiento de la funcionalidad del portal BA Data que garantice los lineamientos establecidos por el organismo.³⁸

³⁸ La DGCIGA respondió en su descargo que “Se trabajará en conjunto con DGARD de manera de generar una estrategia de ejecución de proyectos afectados a la modificación y mejoramiento de la funcionalidad del portal BA Data y que permita contemplar planes de corto, mediano y largo plazo asignados al sistema”

- Entregar, dar servicio y soporte - DSS01
 - Gestionar las operaciones
 - 12. Se encontraron ciertos errores en la funcionalidad del sitio referidos a las versiones anteriores de los datasets o las diferencias de datos entre ellos. La operatividad de las funcionalidades sirve para garantizar la calidad del producto ofrecido.

- Entregar, dar servicio y soporte - DSS02
 - Gestionar las peticiones y los incidentes de servicio
 - 13. Se observó que el sitio web no cuenta con un apartado de “contáctenos”, que tampoco están indicadas claramente las direcciones de email para comunicarse, o los canales alternativos en redes sociales. Explicitar los canales de comunicación facilitan la solución de posibles incidentes y la retroalimentación (*feedback*) por parte de los usuarios del sistema.

VII. RECOMENDACIONES.

Se formulan las siguientes recomendaciones:

- Alinear, Planificar y Organizar - APO02
 - Gestionar la estrategia
 - 1. Es de una importancia notoria la necesidad de la definición de una estrategia de gestión integral de la información pública para el GCABA para que la información generada en las diversas reparticiones se encuentre disponible y pueda ser accedida y utilizada de manera general, compartida, segura y vinculada a otros sistemas. Con ello no solo se pretende alcanzar el objetivo de un gobierno abierto, sino que es el fundamento para la toma de decisiones de políticas públicas con base en la evidencia.

- Alinear, Planificar y Organizar - APO03
 - Gestionar la arquitectura empresarial
 - 2. Se aconseja que se proceda a la definición del universo de la información pública que generan las distintas reparticiones del GCABA, o el establecimiento de criterios o definiciones estratégicos de qué información es más relevante y que luego

sería publicada en el portal de datos abiertos Buenos Aires Data.

3. Se insta a que el flujo de la información sea continuo y automático; que la información que va a ser publicada no deba ser solicitada para cada caso por las áreas responsables de la publicación de la información (DGARD – DGCIGA), y que por ello se vea afectada por desactualizaciones; y que se reduzcan la cantidad de datasets de actualización “eventual”.
4. Es beneficiosa una mayor disponibilidad de soluciones (APIs) que permitan la disponibilidad inmediata de la información a medida que es generada, para una mayor calidad de las prestaciones del sistema.

➤ Alinear, Planificar y Organizar - APO07

- Gestionar los recursos humanos

5. Es sumamente conveniente que los agentes que prestan tareas técnicas dentro de la DGARD tengan su relación laboral formalizada y adecuadamente retribuida, para evitar la pérdida o alta rotación de personal técnico informático.

➤ Alinear, Planificar y Organizar - APO09

- Gestionar los acuerdos de servicio

6. Se sugiere la suscripción de Acuerdos de Niveles de Servicio (SLA)³⁹ por los servicios tecnológicos recibidos. Así como la firma de acuerdos formales sobre condiciones y periodicidad de publicación de la información entre las reparticiones proveedoras de los datos y las encargadas de publicarlos.

➤ Alinear, Planificar y Organizar - APO14

- Gestionar los datos

7. Se recomienda la revisión de los datasets cuyo el plazo de la frecuencia de actualización vencido. Asimismo, de los datos contenidos desactualizados, con errores y omisiones dentro de algunos de los datasets.

³⁹ SLA Service Level Agreement, (en español, Acuerdo de Nivel de Servicio), es un acuerdo firmado entre el proveedor de un servicio y el cliente. El objetivo es establecer el nivel del servicio y sus parámetros, tales como los plazos de entrega, responsables, tiempos de respuesta, tipo de soporte técnico, etc

8. Deberían existir controles automáticos, por medio de softwares, que intenten garantizar que se cumplan los plazos de actualización por parte de las reparticiones; que la información se encuentre efectivamente actualizada, normalizada, sin errores manifiestos o incompleta.
 9. Es beneficiosa la realización, junto con la ASINF, de pruebas de restauración específicas de las VMs que componen el sistema BA Data, como es un buen indicador de operatividad de las políticas de respaldo de la información.
- Construir, Adquirir e Implementar - BAI09
 - Gestionar los activos
 10. Es conveniente el establecimiento de procesos de gestión de activos de información, tales como el propietario del dato, usuario y custodio de la información utilizada; así como procedimientos de evaluación y clasificación de la información según su criticidad, nivel de integridad, confidencialidad, disponibilidad y privacidad.
 - Construir, Adquirir e Implementar - BAI11
 - Gestionar los proyectos
 11. Se sugiere generar planes de corto, mediano y largo plazo para la modificación y mejoramiento del del portal BA Data.
 - Entregar, dar servicio y soporte - DSS01
 - Gestionar las operaciones
 12. Se indica la necesidad de corregir los errores indicados en la funcionalidad del sitio referidos a las versiones anteriores de los datasets o las diferencias de datos entre ellos.
 - Entregar, dar servicio y soporte - DSS02
 - Gestionar las peticiones y los incidentes de servicio
 13. Se observó que el sitio web no cuenta con un apartado de “contáctenos”, que tampoco están indicadas claramente las direcciones de email para comunicarse, o los canales alternativos en redes sociales. Explicitar los canales de comunicación facilitan la solución de posibles incidentes y la

retroalimentación (*feedback*) por parte de los usuarios del sistema.

VIII CONCLUSIONES.

Como primer punto a abordar, es necesario subrayar el importante objetivo de gobierno abierto buscado por el portal y las prestaciones que provee. El Gobierno abierto implica tanto la apertura de los datos públicos como de procesos internos de la administración, basándose en los principios de transparencia, colaboración y participación.

En su gestión diaria, El GCABA genera muchísima información en cada una de sus reparticiones. Pero dicha información suele quedar circunscripta a la repartición, no se halla disponible para ser accedida, compartida o consumida.

Si bien cada repartición es “dueña” del dato (*data owner*), eso no significa que deba utilizarlo -si es que lo hace- para sus fines organizativos propios únicamente. No es correcto que ellas puedan considerar a su información como privativa y vuelvan celosas guardianas de la misma.

En el presente informe se pudo comprobar que la información se encuentra localizada en cada repartición, y para accederla es necesario lograr acuerdos de apertura de datos entre la repartición y las áreas encargadas de la publicación de la información pública en el portal de datos abiertos del GCABA. También se encontró que el flujo de la información no es continuo ni automático y ésta se ve afectada por desactualizaciones, problemas en la normalización y formatos de la información.

En cuanto a la gestión de la información, el GCABA debe considerarse a sí mismo como una sola organización, y debe tratar la información de una manera íntegra, consistente, interrelacionada, compatible y coherente.

No debe entenderse lo anterior como propuesta de un área que sea la única que se encargue de recabar la información de manera centralizada. Sino que se indica como necesario definir políticas y articular soluciones tecnológicas para que la información generada y tratada por una repartición sea accesible (en las condiciones de seguridad y privacidad que el caso demande) y utilizable por otras, y en especial la conducción del poder ejecutivo.

La información producida por un sector del gobierno seguramente sea un insumo de otro y, de lleno, un punto de partida para las decisiones en la gestión que realizan los altos funcionarios.

“Año del 40° Aniversario de la Guerra de Malvinas. En homenaje a los veteranos y caídos en la defensa de las Islas Malvinas y el Atlántico Sur”

Todo lo anterior redundaría en mejorar la experiencia de los ciudadanos y el público en general, ayudar a definir las políticas públicas en base a la evidencia, y lograr la economía, eficacia, eficiencia, equidad y sostenibilidad en las políticas y acciones de gobierno.

ANEXO I⁴⁰

Descripción de Responsabilidades Primarias Dirección General de Calidad Institucional y Gobierno Abierto (DGCIGA)

- Asistir al Subsecretario en la definición e implementación de las políticas de gobierno abierto en articulación con las áreas de gobierno.
- Planificar y ejecutar las políticas de comunicación vinculadas con el desarrollo de un gobierno abierto al ciudadano, en coordinación con las áreas competentes.
- Asistir al Subsecretario en la participación activa en redes nacionales e internacionales de innovación institucional y gobierno abierto para promover el intercambio de buenas prácticas.
- Planificar y ejecutar las acciones de apertura de datos mediante el desarrollo de bases de datos de libre acceso, en coordinación con las áreas competentes y promoviendo la reutilización de datos por parte de la sociedad civil y el sector privado.
- Administrar plataformas de datos abiertos de la Ciudad Autónoma de Buenos Aires, incluyendo BA Data y BA Obras, en coordinación con la Secretaría de Innovación y Transformación Digital de la Jefatura de Gabinete de Ministros.
- Promover acuerdos bilaterales y multilaterales que favorezcan la apertura de datos, en coordinación con las áreas competentes.
- Proponer instancias de colaboración con otros poderes del Estado y otros gobiernos de ciudades y entidades subnacionales de la República Argentina para extender el alcance de las políticas de gobierno abierto.

Gerencia Operativa Articulación Institucional. Acciones:

- Apoyar la coordinación técnica y ejecutiva de ámbitos de relación institucional e interjurisdiccional de las iniciativas de calidad e innovación institucional y de sistema de gobierno, mediante la formulación de propuestas de agenda, la elaboración de minutas de las reuniones, y el seguimiento de los acuerdos alcanzados.
- Identificar riesgos y oportunidades de mejora en la gestión de iniciativas de calidad e innovación institucional y de sistema de

⁴⁰ <https://documentosboletinoficial.buenosaires.gob.ar/publico/PE-DEC-AJG-AJG-463-19-ANX-1.pdf> páginas 20 y 21 [accedido el 09/12/2021]

gobierno, y proponer líneas de acción y mitigación para el logro de las reformas institucionales.

- Realizar estudios, informes y reportes periódicos de seguimiento y monitoreo de las iniciativas de calidad institucional.
- Apoyar a la Dirección General en la implementación de políticas públicas tendientes a mejorar la calidad institucional del Gobierno de la Ciudad Autónoma de Buenos Aires, teniendo como referencia los lineamientos que la Organización para la Cooperación y el Desarrollo Económicos (OCDE) propone para las diversas agendas.
- Apoyar a la Dirección General en la implementación y seguimiento de políticas de Buen Gobierno Corporativo en las Empresas Públicas de la Ciudad Autónoma de Buenos Aires.
- Gestionar la participación de la Ciudad Autónoma de Buenos Aires en los grupos de trabajo sobre Transparencia, Gobierno Abierto y anticorrupción de foros internacionales, garantizando el cumplimiento de los acuerdos establecidos.

Gerencia Operativa Gobierno Abierto. Acciones:

- Asistir en el diseño e implementación de proyectos y políticas de gobierno abierto en el ámbito de la Ciudad Autónoma de Buenos Aires.
- Generar espacios de colaboración con los distintos actores de la ciudadanía.
- Diseñar y promover visualizaciones y productos analíticos con los datos abiertos de la Ciudad, en coordinación con las áreas competentes.
- Asesorar a las reparticiones del Gobierno de la Ciudad Autónoma de Buenos Aires para el mejoramiento de la calidad de la información.
- Generar tableros de control basados en datos de gestión para la toma de decisiones.
- Coordinar con las áreas competentes la realización de talleres y programas para fomentar el conocimiento y la reutilización de datos abiertos por parte de la ciudadanía, especialmente con fines académicos o de desarrollos de negocios.
- Coordinar con la Dirección General Arquitectura de Datos de la Jefatura de Gabinete de Ministros el contacto con los enlaces de datos de cada ministerio y repartición del GCBA a fin de asegurar la publicación en tiempo y forma de los datos abiertos.
- Desarrollar y mantener plataformas basadas en datos abiertos como BA Obras, Presupuesto Abierto, Compras Coronavirus, Ecosistema de Gobierno Abierto y BA Cambio Climático, que permitan a la ciudadanía conocer más sobre temas estratégicos del gobierno de una forma cercana y amigable.

“Año del 40° Aniversario de la Guerra de Malvinas. En homenaje a los veteranos y caídos en la defensa de las Islas Malvinas y el Atlántico Sur”

- Coordinar con la Dirección General Arquitectura de Datos de la Jefatura de Gabinete de Ministros el desarrollo y mantenimiento del portal de datos abiertos BA Data.
- Sistematizar un catálogo de herramientas desarrolladas en código abierto, tendiente a que otras entidades públicas o privadas puedan utilizarlas y/o modificarlas libre y gratuitamente.

ANEXO II

Presupuesto 2020 - Jurisdicción 20 - Programa 54 – Objetivos.

- Articular y ejecutar la estrategia de Gobierno Abierto del GCBA.
- Proponer reformas y/o intervenciones para promover mayor transparencia, rendición de cuentas y participación ciudadana en la gestión de gobierno.
- Coordinar, profundizar y expandir iniciativas transversales orientadas a promover un gobierno abierto y mejorar la calidad de las instituciones públicas.
- Llevar adelante investigaciones, seminarios y capacitaciones para abordar temas vinculados con la transparencia, rendición de cuentas e innovación institucional.
- Promover y garantizar el cumplimiento de estándares internacionales en materia de transparencia institucional, gobierno abierto e integridad.
- Coordinar con el gobierno federal y los gobiernos subnacionales iniciativas en materia de transparencia, calidad institucional, gobierno abierto e innovación gubernamental.
- Coordinar iniciativas transversales para el cumplimiento de los objetivos del Segundo Plan de Acción de Gobierno Abierto de la Ciudad, cocreación del Tercer Plan de Acción de Gobierno Abierto de la Ciudad, y la sistematización y posicionamiento de buenas prácticas del GCBA en gobierno y datos abiertos.
- Coordinación de la Política de Datos Abierto del GCBA.
- Diseñar y ejecutar iniciativas vinculadas a la apertura y reutilización de datos sectoriales (educación, salud, género, transporte, etc.) en el marco de una coordinación transversal de las políticas de datos abiertos.
- Diseñar y ejecutar nuevas iniciativas que promuevan la participación, cocreación y colaboración ciudadana.
- Mantener y actualizar los portales y repositorios Github correspondientes a las iniciativas BA Data y BA Obras.

ANEXO III⁴¹

Descripción de Responsabilidades Primarias Dirección General de Arquitectura de Datos (DGARD)

- Definir estándares para asegurar la calidad e interoperabilidad de los datos.
- Establecer los criterios para niveles de acceso y seguridad de los datos, en coordinación con la Secretaría Legal y Técnica.
- Implementar y administrar el Inventario de Metadatos del Gobierno de la Ciudad Autónoma de Buenos Aires.
- Mantener y administrar el portal de datos de la Ciudad Autónoma de Buenos Aires, en coordinación con la Secretaría General y Relaciones Internacionales.
- Realizar el procesamiento y control de los datasets, disponibilizándolos para su publicación en el portal de datos abiertos, y proveer asesoramiento a las áreas del Gobierno de la Ciudad Autónoma de Buenos Aires para mejorar la calidad de la información, en coordinación con la Secretaría General y Relaciones Internacionales.
- Implementar y administrar la infraestructura de datos espaciales de la Ciudad Autónoma de Buenos Aires.
- Definir el lineamiento y el mantenimiento de las entidades interoperables.
- Definir e implementar la infraestructura tecnológica de la Oficina Central de Datos.
- Coordinar la provisión de la infraestructura necesaria con la Agencia de Sistemas de la Información.
- Desarrollar y mantener el repositorio de datos del Gobierno de la Ciudad Autónoma de Buenos Aires
- Asistir a las áreas de Gobierno, en la incorporación de herramientas informáticas de análisis y administración de datos, que permitan generar información para el seguimiento, control y análisis de sus actividades, la creación de modelos predictivos y la toma de decisiones.

Gerencia Operativa Estandarización e Interoperabilidad de datos (GOEI)

⁴¹ <https://documentosboletinoficial.buenosaires.gob.ar/publico/PE-DEC-AJG-AJG-282-21-ANX-1.pdf> [accedido el 09/12/2021]

- Coordinar y mantener las entidades interoperables del Gobierno de la Ciudad Autónoma de Buenos Aires.
- Establecer el protocolo de intercambio de datos y brindar soporte a las áreas que lo requieran, en coordinación con las áreas competentes.
- Asistir a las áreas de gobierno en la normalización, consistencia y coherencia de sus datos. Promover nuevas regulaciones en la administración y gestión de los datos y definir su obligatoriedad.
- Definir protocolos y guías para la correcta anonimización de la información.

Gerencia Operativa Inventario de Metadatos (GOIM)

- Mantener y actualizar el Inventario de metadatos del Gobierno de la Ciudad Autónoma de Buenos Aires.
- Brindar asesoramiento y soporte a las áreas de gobierno para la utilización y actualización de sus metadatos.
- Implementar el lineamiento de clasificación de datos.
- Mantener y administrar el portal de datos abiertos de la Ciudad Autónoma de Buenos Aires, en coordinación con las áreas competentes.
- Asistir a la Dirección General de Arquitectura de Datos en el mantenimiento y administración de la Infraestructura de Datos Espaciales de la Ciudad Autónoma de Buenos Aires.
- Colaborar con la Dirección General en el procesamiento y control de los datasets, disponibilizándolos para su publicación en el portal de datos abiertos.

Gerencia Operativa Ingeniería de Datos (GOLD)

- Elaborar, difundir y actualizar los lineamientos para el diseño de la arquitectura de las aplicaciones, bases de datos y la integración de la información, en coordinación con las áreas competentes.
- Diseñar el modelado de los datos y desarrollo de la arquitectura para la integración de información.
- Desarrollar y mantener procesos automáticos de integración de datos con el objetivo de facilitar la explotación de los mismos.
- Definir herramientas para la integración y modelado de los datos.
- Disponibilizar herramientas para la recolección de datos en campo que se integren a los modelos de datos.
- Diseñar, elaborar y mantener herramientas geográficas, con base en el mapa interactivo de la Ciudad. Asistir a la Dirección General de Arquitectura de Datos en la implementación y mantenimiento de la infraestructura tecnológica de la Subsecretaría de Políticas Públicas Basadas en Evidencia.
- Asistir a la Dirección General de Arquitectura de Datos en la implementación de la infraestructura de datos espaciales de la Ciudad Autónoma de Buenos Aires.

ANEXO IV

Presupuesto 2020 - Jurisdicción 21 - Programa 109 – Objetivos.

- La Dirección General de Ciencias de la Información es responsable de diseñar y coordinar programas y proyectos de innovación que permitan la transformación y mejora de los servicios de Gobierno a través de métodos de ciencia de datos, análisis y visualización, generación y mantenimiento de mapas digitales e innovación estratégica. Asimismo, elaborar tableros de control en coordinación con las áreas competentes, que faciliten el proceso de toma de decisiones basadas en datos.
- El Programa tiene como principal objetivo brindar herramientas y servicios que permitan a las áreas de Gobierno tomar decisiones eficientes en base a evidencias. En este marco se proveen servicios de integración, normalización, inteligencia de datos y tableros de control a las distintas reparticiones de Gobierno, para la toma de decisiones asentadas en información.
- De la misma forma, se trabaja coordinando y administrando las bases de datos georreferenciadas existentes en las áreas del Gobierno de la Ciudad Autónoma de Buenos Aires, confeccionando diferentes servicios, como por ejemplo mapas digitales de la Ciudad.
- En particular, se brindan servicios a las áreas del Gobierno aplicando procesos y técnicas de ciencia de datos, modelos predictivos, análisis de Big Data y aprendizaje automático con el fin de optimizar recursos y automatizar procesos.