

**INTER-
CONEXIÓN
LATAM**

**INTER-
CONEXIÓN
LATAM**

**INFORME
DIAGNÓSTICO
INTERCONEXIÓN
LATAM**

CÓMO CONECTAR LOS DATOS DE AMÉRICA LATINA

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum

Febrero 24 2020

Muchos autores

ENFOQUE DE LA INVESTIGACIÓN

Esta investigación se llevó a cabo por parte de PODER, Codeando México y DataSketch. Si bien fue importante plasmar las particularidades del uso de datos abiertos por parte de organizaciones en Latinoamérica, se tomó la decisión de enfocar los esfuerzos en una muestra representativa, tanto para la encuesta como para las entrevistas, ameritó el uso de metodologías de investigación más flexibles y cualitativas, guiadas por los siguientes enfoques:

- Práctico —basado en lo que las organizaciones han experimentado y realizado en el tema de datos abiertos, más que en únicamente sus percepciones.
- Exploratorio —minimizar al extremo suposiciones y mantener apertura a perspectivas e interpretaciones que no se habían contemplado anteriormente.
- Descriptivo —proporcionar un trabajo de base para futuras investigaciones y/o exploraciones a través de la documentación de la experiencia y desafíos de las organizaciones al momento de trabajar con datos abiertos.

METODOLOGÍA DE INVESTIGACIÓN

2.1 Encuesta

Se distribuyó una encuesta entre organizaciones de América Latina que utilizan y/o generan datos como parte de su trabajo. La encuesta se elaboró en un Google Form y fue compartida entre las redes a las que pertenece cada una de las organizaciones que integran el equipo de trabajo de Interconexión, obteniendo 21 respuestas.

La encuesta se realizó con base a la experiencia y comentarios obtenidos en el primer ejercicio del proyecto, mismo que se llevó a cabo durante AbreLatam 2019 en Quito, y fue de naturaleza exploratoria. Adicional a contestar preguntas sobre la organización, sus desafíos y experiencias al usar datos abiertos, los resultados de la encuesta contribuyeron a la selección final de sujetos para entrevista.

2.2 Entrevistas

Se realizaron entrevistas a profundidad a 7 organizaciones que respondieron la encuesta. Fueron seleccionadas de acuerdo a su interés por participar en el estudio y con el objetivo de que ilustraran un rango representativo de experiencias y en áreas geográficas de la región.

2.3 Análisis

El equipo de Interconexión realizó un mapeo y análisis de la información recolectada a través de la encuesta y casos de estudio; priorizaron, perfilaron, evaluaron e hicieron conexiones entre diferentes variables. Las entrevistas se transcribieron, las frases de los entrevistados se agruparon por temática en un diagrama de afinidad y se realizó un ejercicio para categorizar a las organizaciones en diferentes perfiles. Los aspectos más representativos fueron integrados y desarrollados en este informe en la sección de Diagnóstico.

2.4 Organizaciones Encuestadas

Listado orgs y perfiles

2.5 Limitantes de la investigación

- Perspectiva del equipo de trabajo: las organizaciones involucradas en el proyecto se encuentran involucradas en el ecosistema de datos abiertos y también respondieron la

encuesta. Si bien esta experiencia contribuye a generar ideas útiles para la interpretación de la información recabada, se tomó cuidado de identificar y cuestionar suposiciones que vinieran de una experiencia personal.

- Limitantes inherentes al proceso de investigación: el éxito de la encuestas, especialmente las encuestas en línea, suele depender de qué tanta motivación y comunicación exista. Tanto la encuesta como los casos de estudio se basan mucho en el conocimiento de la organización y el ecosistema. La priorización de actividades programáticas sobre las actividades comunitarias contribuyó a que sean pocas organizaciones las que se tomen el tiempo para participar en este tipo de ejercicios. La encuesta se elaboró con la intención de ser lo más concreta, sencilla y rápida posible, facilitando de esta forma su llenado por parte de las organizaciones.
- Recursos limitados: los recursos disponibles se comprometían en una serie de formas que limitaban el alcance de este análisis:
 - Tiempo limitado para extender los alcances de la investigación
 - Tiempo limitado de un equipo de trabajo reducido

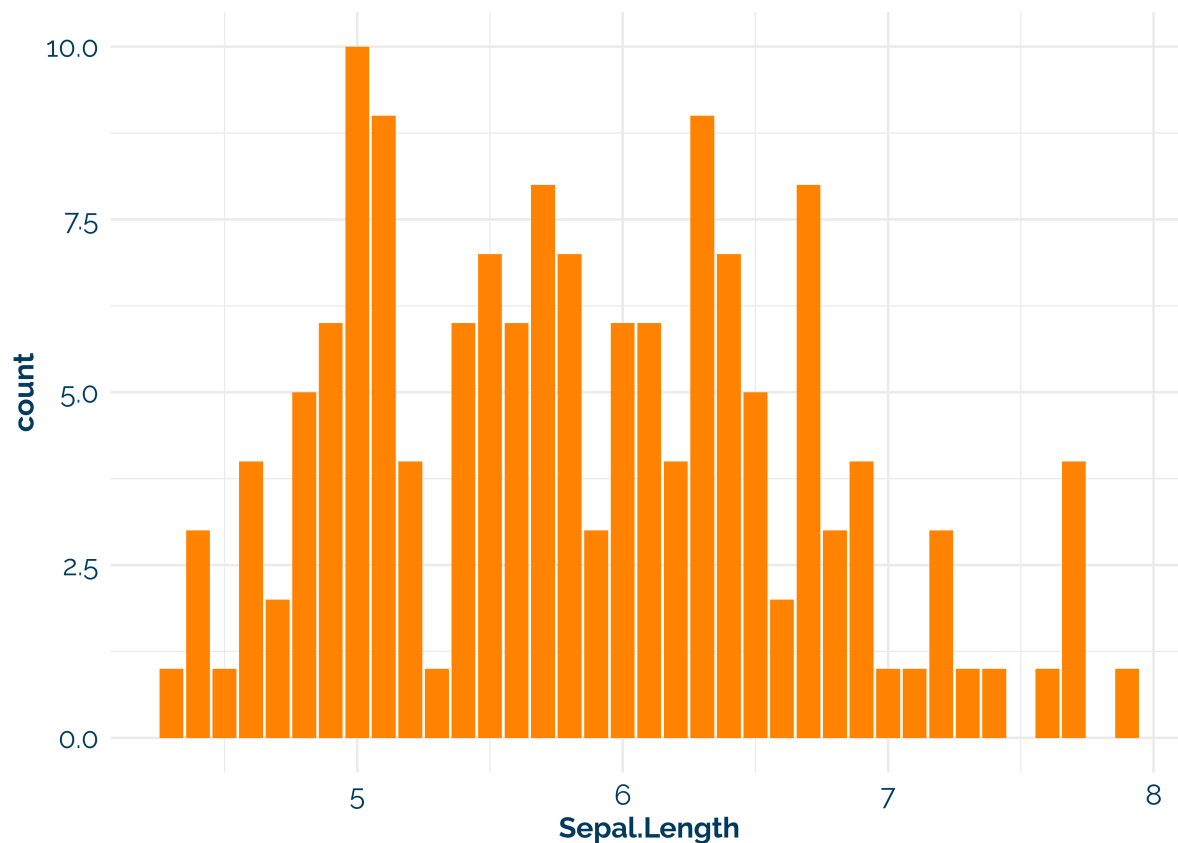
DIAGNÓSTICO

3.1 Uso de datos por parte de organizaciones de sociedad civil en Latinoamérica

¿Qué papel tienen los datos en la estructura de la organización? ¿Cómo describimos la experiencia de las organizaciones usando y generando datos? ¿De dónde se obtienen estos datos? Estas son algunas de las preguntas que se buscó responder a través de las encuestas y casos de estudio realizados con las organizaciones participantes en este primer ejercicio piloto.

En cuanto a la experiencia de las organizaciones en el uso de datos, destacaron el uso de los mismos para investigaciones periodísticas (periodismo de datos), para proyectos de tecnología cívica enfocados a participación ciudadana, transparencia y rendición de cuentas, así como para proyectos de política pública. De igual forma, varias organizaciones mencionaron trabajar específicamente en limpieza y publicación de datos, tanto de otras organizaciones como de instituciones gubernamentales.

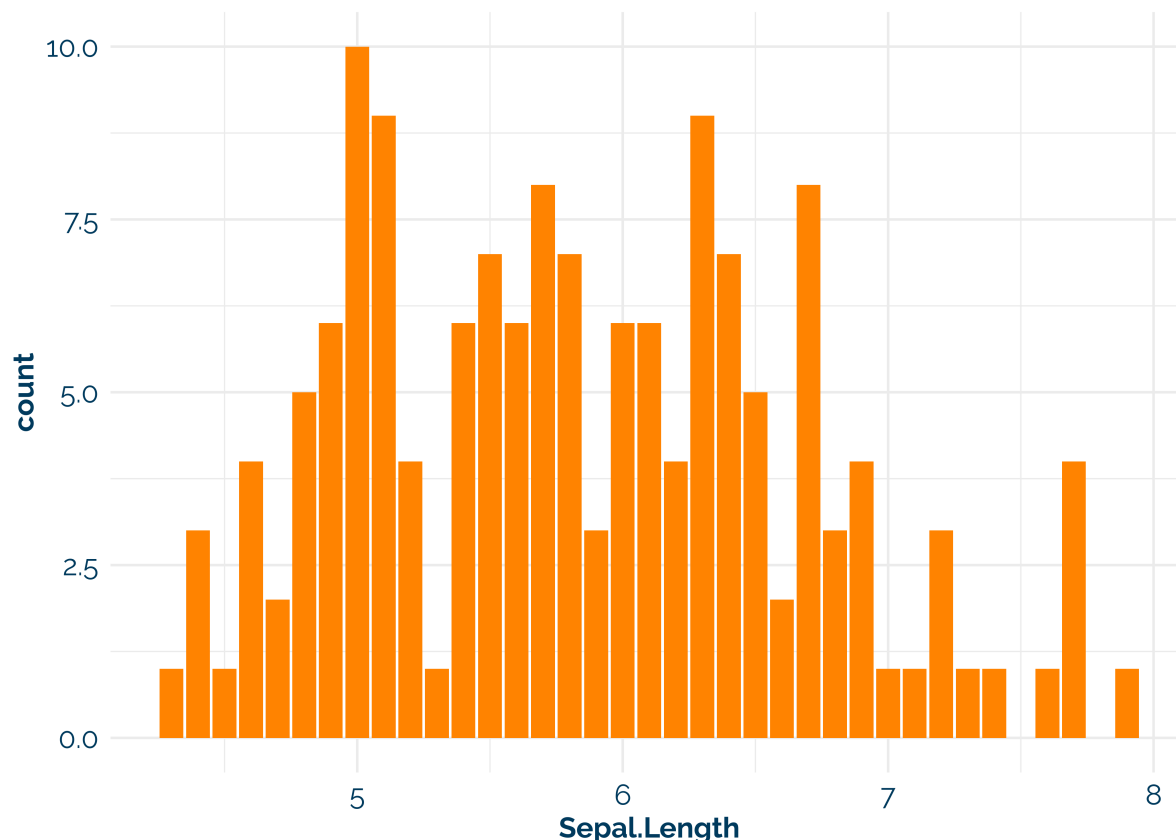
Lo anterior nos lleva al Gráfico 1, donde se puede encontrar de manera más detallada las diferentes formas en las que las organizaciones participantes utilizan datos habitualmente, siendo las más destacadas consumo de bases de datos como archivos planos (csv, xlsx, json), transformación de datos públicos para hacerlos más útiles, generación de informes, tableros, visualizaciones para incidencia o toma de decisiones. Esto cobra un especial sentido si entendemos que las organizaciones obtienen datos que después transforman en información digerible y/o investigaciones periodísticas que presentar a sus audiencias.



Entendiendo el uso que le dan las organizaciones a las diferentes bases de datos a las que acceden, la pregunta siguiente sería de dónde se obtienen esas bases de datos, a lo cual las organizaciones enfatizaron que a través de portales gubernamentales de datos abiertos e instituciones descentralizadas, en diversas comunidades y de otras organizaciones de sociedad civil y mayoritariamente de solicitudes de información a diferentes organismos gubernamentales.

En relación a los datos que obtienen de otras comunidades u organizaciones de sociedad civil, fueron pocas las organizaciones que dijeron hacer uso constante (más de 5 veces) de datos de otras organizaciones, únicamente un 20.8% (Gráfico 2). La mayoría mencionó a lo largo de la encuesta nunca haber utilizado datos de otras organizaciones en sus proyectos (37.5%).





Algunas de las razones mencionadas por las organizaciones por las que no utilizan datos de otras organizaciones son:

Los datos se obtuvieron de fuentes oficiales Existe una falta de acceso a los datos, ya sea porque no se publican o porque no hay voluntad para compartirlos. No existen los datos

Volteando la mirada hacia la generación de datos en la región, a lo largo de la encuesta compartida, varias organizaciones mencionaron haber participado en procesos de apertura de datos. Esto da cabida a, en caso de realizarse una investigación a mayor profundidad, entender si los ejercicios de generación de datos son esfuerzos inherentes a la organización o colaboraciones transversales. pero no queda claro en sus respuestas si ellas generaron esos datos y si es algo que hacen constantemente dentro de la organización o sólo como ejercicios aislados en los que participaron.

Relativo a las solicitudes de información obtenidas, se recalcó el trabajo que suelen realizar para generar bases de datos a partir de la información contenida en dichas solicitudes, para posteriormente publicar esas bases.

Varias organizaciones mencionaron que el trabajo que hacen en la generación de bases de datos no se publica constantemente, ya que siguen siendo procesos muy artesanales y no cuentan con los conocimientos técnicos ni infraestructura para poder ello. Adicionalmente, se generan bases de datos a partir de las investigaciones que se realizan, pero de nuevo se establece la falta de un formato único y accesible para que estos datos realmente estén abiertos. Sin embargo, se mencionó en varias ocasiones durante las entrevistas que los productos que generan las organizaciones derivados de estos datos siempre son publicados con licencias y/o códigos abiertos.



Aquellos datos que sí se publican, ¿cómo son publicados?, ¿se utiliza algún estándar específico por parte de las organizaciones? Este tema se abordó de forma específica en la encuesta, solicitando a las organizaciones que mencionaran los estándares de datos abiertos con los que trabajan. Como se puede observar en el Gráfico 3, el 33.3% de las organizaciones nunca han usado un estándar para la publicación de datos abiertos. En cuanto a aquellas que sí utilizan alguno, el Open Data Contracting Standard (ODCS) y Popolo fueron los más mencionados. Adicionalmente, se mencionaron estándares como Akoma Ntoso, Beneficial Ownership Data Standard, entre otros.

Gráfico 3

La organización Observatorio del Poder hizo particular énfasis en la falta de estándares tanto en el formato abierto de los datos, como en cuál es la información que por ley se debe compartir, así como la necesidad de homologar ambos aspectos a nivel regional para que sea posible compartir la información.

En cuanto al valor de los datos y su impacto en la estructura de la organización, la mayoría de las organizaciones establecen los datos abiertos como su eje central, ya sea para investigaciones periodísticas o para proyectos de tecnología cívica. Los datos abiertos son la espina dorsal de la organización de acuerdo a la encuesta y casos de estudio realizados.

No obstante, algunas organizaciones mencionan a los datos no como un eje rector, pero sí como un elemento de especial relevancia dentro de la estructura y en el trabajo de la organización. Estas organizaciones mencionaron tener un área enfocada a datos abiertos, sin especificar las actividades específicas que realizan en estas áreas.

De acuerdo a los casos de estudio analizados, el uso de datos le ha dado valor agregado y metodologías más robustas al trabajo que ya están haciendo las organizaciones, ya sea en temas como periodismo de datos como en diversos proyectos que manejan. Esto se complementa con lo analizado en la encuesta, donde varias organizaciones mencionaron a los datos como una herramienta que les ha dado método y estructura a la hora de diseñar y desarrollar sus procesos y proyectos.

La organización Ojoconmipisto recalcó en su entrevista que los datos han permitido explorar nuevas formas de socializar y presentar la información para hacerla accesible al público y tener un mayor impacto en sus proyectos, algo que se repitió entre el resto de las organizaciones entrevistadas.

Para trabajar con bases de datos abiertos es necesario contar con algunas habilidades técnicas, algo que se buscó abordar tanto en la encuesta como en las entrevistas con las organizaciones, con el objetivo de entender tanto la capacidad del personal de la organización como posibles limitantes que se encontraran en este aspecto.

Todas las organizaciones hicieron énfasis en que sus equipos de trabajo aprenden o desarrollan una serie de habilidades complementarias para ser capaces de realizar el análisis de las bases de datos que obtienen. De igual forma, dichos análisis también se complementan con procesos más artesanales como herramientas hemerográficas, especialmente en temas donde hay poca información y datos.

Desafíos en el uso de los datos



Uno de los desafíos mencionado en repetidas ocasiones a lo largo de la encuesta y las entrevistas, y del cual se ha hecho mención brevemente en el apartado anterior, es la falta de formatos y/o estándares para la generación y publicación de datos abiertos, siendo aún un trabajo muy artesanal y que se produce sin rigor, desde el entendimiento y contexto de cada organización. Al no contar con una homologación de estándares, es muy difícil que bases de datos sobre los mismos temas puedan ser analizadas a un nivel macro para detectar patrones y coincidencias. Los estándares no sólo tienen que homologarse entre organizaciones que están trabajando un proyecto específico, sino en toda la región latinoamericana.

Adicional a la falta de estándares, se preguntó a las organizaciones cuáles eran las barreras más recurrentes a las que se enfrentaban a la hora de compartir datos, siendo la actualización y la verificabilidad las opciones que más se mencionaron en la encuesta después de la falta de formatos (Gráfico 4).

Gráfico 4

En las entrevistas se habló en repetidas ocasiones de una falta de acceso a la información e incluso una falta de información en sí misma. La falta de estándares no sólo afecta la capacidad de generar y publicar información por parte de las organizaciones, también genera vacíos legales que las instituciones gubernamentales utilizan para obstaculizar el acceso a cierta información.

Los esfuerzos por parte de los gobiernos para aperturar datos y garantizar el acceso a la información aún son muy recientes en la mayoría de los países latinoamericanos, por lo que no existen estrategias claras y robustas, tanto desde el gobierno como desde la sociedad civil. Este problema también se deriva de la falta de voluntad política para liberar la información. Existe una colaboración reactiva por parte de los gobiernos para sólo publicar la información requerida por ley, pero hay muy pocos esfuerzos reales por aperturar la información en automático, de manera proactiva. El discurso de apertura no se traduce en acciones concretas por parte de los gobiernos y esto se confirma al ver que las solicitudes de información siguen siendo una herramienta de transparencia y acceso a la información muy recurrente entre las organizaciones.

Asimismo, las instituciones gubernamentales tampoco cuentan con conocimientos sobre estándares adecuados y buenas prácticas para generar datos, lo cual se traduce en impactos negativos particularmente en organizaciones que trabajan temas de género o inclusión de la comunidad LGBT+. La organización Luchadoras hizo énfasis en lo anterior al mencionar que las instituciones gubernamentales han fallado en recopilar información de manera segregada, lo que crea un vacío de datos relacionados a temas de género en México. De igual forma, Visibles, organización enfocada en temas de salud para la comunidad LGBT+, habló durante la entrevista de una falta de datos en estos temas, siendo las organizaciones quienes han hecho mucho del trabajo de generar esta información, pero no de una manera homologada.

Otro de los desafíos mencionados particularmente por las dos organizaciones mencionadas en el párrafo anterior hace referencia a la complejidad de la anonimización de los datos y la falta de estándares regionales específicamente en este ámbito. Al haber no haberlos, mucha de la información generada no puede compartirse, perpetuando la inexistencia de ciertos datos.

Es importante tomar en cuenta también los desafíos organizacionales, como la falta de recursos económicos y humanos, situación que se mencionó constantemente a lo largo de las entrevistas. El manejo de bases de datos macro implica un desarrollo de habilidades técnicas

que pocas organizaciones tienen y a las que pocas pueden acceder. Esto deriva en que la capacidad y alcance de las organizaciones sea vea reducido, al emplear métodos más artesanales para analizar la información y con cada integrante del equipo realizando una serie de funciones diversas para cubrir las necesidades de la organización.

Por último punto y sobre esta misma línea, se habló también de una falta de infraestructura tecnológica dentro de las organizaciones. En algunos casos, ellas han desarrollado sus propias herramientas de análisis y visualización de datos, pero son pocas las que pueden desarrollar este tipo de productos.

COLABORACIÓN REGIONAL

Además de entender el uso de los datos abiertos en organizaciones de la región latinoamericana, fue de suma importancia conocer el estado de la colaboración y la percepción que tienen dichas organizaciones sobre el mismo, así como saber si es un aspecto relevante a fortalecer en un futuro cercano.

Varias organizaciones, especialmente aquellas enfocadas en periodismo de datos, afirmaron en sus entrevistas que ya existen o se están formando iniciativas y redes entre organizaciones que utilizan datos abiertos para sus investigaciones. Otras dimensionan alguna forma de colaboración en el uso y reuso de información producida por otras organizaciones, algo que se detalló más a fondo en la sección “Uso de datos por parte de organizaciones de sociedad civil en Latinoamérica”.

¿Cómo se lleva a cabo la colaboración en el uso y compartir de datos abiertos en la región?
¿Cuál es el valor que perciben las organizaciones en cuanto a esta colaboración? Por un lado, tenemos organizaciones que afirman conocer esfuerzos colectivos para que exista una colaboración más robusta y se puedan generar y compartir datos de manera más homologada. Por otro lado nos encontramos con organizaciones que perciben la colaboración como algo muy informal, nacida desde las redes personales y el “amiguismo”, pero sin que esto se traduzca necesariamente en la creación de proyectos e iniciativas colectivas.

A pesar de que los esfuerzos por formalizar colaboraciones regionales aún sea puedan concebir como pocos y desarticulados, las organizaciones entrevistadas hicieron particular énfasis en el valor de la colaboración para detectar redes y/o patrones de corrupción regionales en distintos temas. Por ejemplo, la patrones de lavado de dinero y evasión de impuestos detectados en la investigación Panamá Papers, el análisis del estudio de la leche en Perú y otros países de la región, el cual se llevó a cabo con la red Palta.

Internamente, las colaboraciones entre organizaciones contribuyen al desarrollo de nuevas capacidades dentro del equipo, así como a maximizar el alcance de investigaciones o brindar visibilidad a temas que son relevantes, pero se trabajan de manera aislada por diferentes organizaciones.

4.1 Desafíos para una mayor colaboración regional

En el apartado anterior se mencionó la falta de esfuerzos articulados para publicar datos, con pocas y débiles estrategias colectivas para impulsar la agenda de datos abiertos y de apertura de información. La percepción general es que los esfuerzos por impulsar esta agenda se encuentran muy desarticulados, no se vislumbra una línea clara de hacia dónde moverse colectivamente en la región.

Las organizaciones reconocen, sin embargo, que establecer los canales y plataformas adecuados para colaborar puede ser muy complejo, ya que en su mayoría se encuentran al tope de tiempo y recursos. Si bien una vez que un proyecto colaborativo se define y avanza, ya hay un mayor compromiso para trabajar de manera conjunta, pero el proceso de construir y diseñar el proyecto puede tomar mucho tiempo. La sobrecarga de trabajo y la falta de recursos no contribuyen al desarrollo de iniciativas de colaboración, tardando aún más en concretarse.

Otro de los desafíos mencionados en cuanto a la colaboración entre organizaciones fue una persistente cultura de resistencia y recelo al compartir datos, algo que también puede verse en la Gráfica 4 sobre las barreras para compartir la información. Esta falta de voluntad para compartir datos da como consecuencia una falta de confianza en el ecosistema, desincentivando los pocos esfuerzos para una colaboración integral y fortalecida en Latinoamérica.

RECOMENDACIONES

A lo largo de este informe se ha presentado el contexto y las complejidades a las que se enfrentan las organizaciones que utilizan datos abiertos en diferentes países de América Latina. Ello presenta un panorama general, no sólo sobre el uso de datos, sino también relacionado a la colaboración regional y las limitantes a la hora de fortalecerla. Interconexión es una propuesta que busca articular los esfuerzos ya existentes, pero aislados, en un espacio digital que funcione como pilar para la generación y publicación de datos, así como para la co-creación y el desarrollo de proyectos colaborativos entre las organizaciones participantes.

Para lograrlo, se definen una serie de recomendaciones que, desde la visión de las organizaciones entrevistadas y aquellas que componen al equipo de trabajo de Interconexión, tiene la intención de definir estrategias más claras para empujar colectivamente una mayor articulación a nivel regional. Al tener una mayor claridad en los objetivos de estos esfuerzos, se espera ver un compromiso mayor para trabajar de manera homologada. Las alianzas y esfuerzos de colaboración ya existentes son un excelente punto de partida que no debe hacerse a un lado, al contrario, partir de estos esfuerzos ayudará a pasar de la colaboración al co-desarrollo de información.

Es importante entender en qué punto se encuentra cada organización y a partir de ello, identificar las convergencias para ir homologando los esfuerzos para la apertura de datos. Documentar casos de éxito en los que se han generado y compartido datos entre organizaciones o conectado dos herramientas para explicar las historias son recursos que pueden visibilizar de manera más tangible mejores prácticas que puedan replicarse en la región.

La generación de estándares y formatos mínimos que respondan a las necesidades de la región y, más importante, a las necesidades de las organizaciones que los utilizan es un punto clave. Para ello, es importante construir espacios de discusión en donde estas organizaciones pongan sobre la mesa los requisitos necesarios para generar protocolos confiables de anonimización, seguridad, corresponsabilidad y transparencia en los procesos de recopilación y uso de los datos.

En cuanto a la creación de una plataforma digital, esta se debe diseñar desde las necesidades que tiene el personal operativo de las organizaciones, es decir, las personas que manejan y buscan esos datos. La plataforma debe contar con una serie de recursos, no sólo técnicos, sino guías de buenas prácticas, casos de éxito y conocimientos compartidos por otras organizaciones.

Se habla poco de un análisis o un tableros de visualización a lo largo de las encuestas y entrevistas. Si bien se menciona que las herramientas que se llegan a generar se comparten con un código abierto, el reuso del código es un campo a explorar con mayor profundidad.

