

DEFINICIÓN DE LA ESTRATEGIA DE BIG DATA PARA EL ESTADO COLOMBIANO Y PARA EL DESARROLLO DE LA INDUSTRIA DE BIG DATA EN COLOMBIA

*Buenas Prácticas sobre
cómo los Gobiernos pueden
implementar una estrategia
para estimular el ecosistema
de negocios Big Data*

2017-2018

2



DNP Departamento
Nacional
de Planeación



Producido por un equipo compuesto por:

Viviana Cañón, Andrés Clavijo, Luis Godoy,
Emmanuel Letouzé, Gabriel Pestre, Julie Ricard

Coordinado por Emmanuel Letouzé

Con insumos de David Shrier

Bajo la supervisión general de: Alex Pentland

Diagramación editorial por Paola Caile

Asistencia de investigación y edición por Natalie
Grover y Rodrigo Lara Molina

Noviembre de 2017

Versión revisada y ajustada

**BUENAS PRÁCTICAS SOBRE
CÓMO LOS GOBIERNOS
PUEDEN IMPLEMENTAR UNA
ESTRATEGIA PARA ESTIMULAR
EL ECOSISTEMA DE NEGOCIOS
BIG DATA**



ÍNDICE

Glosario	6
A. Introducción:	8
1. Alcance	8
2. Metodología	10
B. Áreas de acción clave	10
1. Gobernanza de data crumbs: el balance entre generar confianza y transparencia y promover el uso de Big Data	10
1.1. Protección de datos: la clave de un sistema cibernético seguro	10
1.2. Garantizar la privacidad y promover prácticas éticas	14
1.2.1 Garantizar la privacidad de los ciudadanos: el ejemplo del GDPR	14
1.2.2 Promover prácticas éticas para orientar el uso de Big Data por el sector privado	16
1.3. Datos abiertos para impulsar la economía	16
1.3.1 Promover una mayor apertura de los datos públicos, a diferentes niveles del gobierno	17
1.3.2 Promover o regular la apertura de datos en otros sectores (gobiernos locales, academia, o empresas)	17
1.3.3 Impulsar el “emprendimiento de datos”, a través del apoyo a la innovación y la incubación de nuevas empresas	18
2. Construir y asegurar capacidades humanas y de capital	19
2.1. Mejorar la educación y las competencias en análisis de datos	19
2.1.1 Una formación en datos y TIC en etapa escolar	20
2.1.2 Una formación en datos y TIC en educación superior, en el mundo profesional y ciudadano	20

2.2.	Investigación y Desarrollo para Big Data	23
2.3.	Modernizar la infraestructura de datos	25
3.	Construyendo comunidad en torno al Big Data	28
3.1.	Facilitar Alianzas Público-Privadas	28
3.2.	Facilitar el desarrollo de clusters a nivel local	34
3.2.1	Facilitar el desarrollo de “Big Data hubs”	34
3.2.2	Empoderar a los ciudadanos para consolidar un ecosistema de Big Data seguro y justo	35
3.2.3	Lanzar iniciativas públicas que conecten a los actores del ecosistema	35
	Observaciones finales	36
	Bibliografía	37

GLOSARIO

El presente glosario se construye sobre algunos de los conceptos incluidos en el glosario del documento anterior, estado del arte y análisis comparativo de estrategias nacionales de Big Data. Este glosario incluye conceptos nuevos que son centrales para el desarrollo del documento.

Analítica de big data [Fuente: IBM]

Es el uso de técnicas analíticas avanzadas contra conjuntos de datos muy grandes y diversos que incluyen datos estructurados, semiestructurados y no estructurados, de diferentes fuentes y en diferentes tamaños, desde terabytes hasta zettabytes. El análisis de Big Data permite a los analistas, investigadores y usuarios comerciales tomar decisiones mejores y más rápidas utilizando datos que antes eran inaccesibles o inutilizables. Mediante el uso de técnicas de análisis avanzadas como análisis de texto, aprendizaje de máquinas, análisis predictivo, minería de datos, estadísticas y procesamiento de lenguaje natural, las organizaciones pueden analizar fuentes de datos antes no explotados, de forma independiente o junto con sus datos empresariales existentes, para obtener nuevas ideas que les permitan tomar decisiones mejores y más rápidas.

Computación en la nube [Fuente: Amazon Web Services]

La computación en la nube es la entrega a la demanda de capacidad de cómputo, almacenamiento de bases de datos, aplicaciones y otros recursos de TI, mediante una plataforma de servicios en la nube (a través de Internet) con precios de pago por uso. La plataforma de servicios en la nube brinda acceso rápido a recursos de TI flexibles y de bajo costo. Evita inversiones iniciales en hardware y administración de este y en cambio, permite externalizar dicho servicio y ajustar el tipo y el tamaño correctos de los recursos informáticos necesarios con mayor facilidad.

Distribución de Hadoop

En el contexto de software libre, una distribución es una colección de componentes de software diseñados, ensamblados y configurados para que puedan ser utilizados “tal como están” para su propósito previsto.

Hadoop es un software Apache, diseñado para manejar y explotar big data. Está disponible para su descarga y utilización, pero las distribuciones de este, ofrecen una alternativa más fácil de usar, con versiones gratuitas y pagadas. Esto es análogo a lo que sucede con Linux y sus distribuciones como RedHat, Suse y Ubuntu. Algunas distribuciones de Hadoop son: Cloudera, Hortonworks, MapR, IBM e Intel.

Enterprise data warehouse (almacén de datos) [Fuente: Sonra]

Son repositorios centrales de datos integrados y depurados de una o más fuentes diferentes. Almacenan datos actuales e históricos. Está diseñado para almacenar datos estructurados de sistemas transaccionales, es un sistema utilizado para la generación de informes y el análisis de datos, y se considera un componente central de la inteligencia de negocios. Se compone de bases de datos relacionales y en el centro de su funcionamiento están los procesos de extracción, transformación y carga (ETL en inglés).

Economía digital [Fuente: OECD]

La amalgama de varias tecnologías de uso general y la gama de actividades económicas y sociales que realizan las personas a través de Internet y las tecnologías relacionadas. Abarca la infraestructura física en la que se basan las tecnologías digitales (líneas de banda ancha, enrutadores), los dispositivos que se utilizan para acceder (computadoras, teléfonos inteligentes), las aplicaciones que proporcionan (Google, Facebook) y la funcionalidad que proporcionan (internet de las cosas, análisis de datos , computación en la nube).

Emprendimiento de datos abiertos [Fuente: Open Data Incubator Europe]

Emprendimientos de pequeño y mediano tamaño que hacen parte del ecosistema de Datos Abiertos y explotan, generan o facilitan la generación de datos abiertos para construir negocios sostenibles que generen impacto económico, social y/o ambiental.

Filtración de datos personales [Fuente: Privacy Rights Clearinghouse]

Es la divulgación intencional o no intencional de información segura o privada/confidencial a un entorno que no es de confianza. Los incidentes van desde ataques concertados por hackers asociados con el crimen organizado, activismo político o gobiernos nacionales, hasta la eliminación descuidada de equipos informáticos usados o medios de almacenamiento de datos. Las infracciones de datos pueden implicar información financiera, como datos de tarjetas de crédito o bancarias, información personal de salud, información de identificación personal, secretos comerciales de corporaciones o propiedad intelectual. La mayoría de las filtraciones de datos involucran datos no estructurados, sobreexposados y vulnerables: archivos, documentos e información confidencial. Privacy Rights Clearinghouse afirma que desde 2005 hasta la fecha se han hecho públicas 7.730 filtraciones de datos.

Industria Big Data [Fuente: Alex Pentland]

El conjunto de actores y actividades que generan ingresos a partir de datos, en uno o más puntos dentro de la cadena de valor de datos; incluye seguridad, servicios, infraestructura de almacenamiento, redes, infraestructura de centros de datos, herramientas de descubrimiento, aplicaciones de big data, análisis, etc. También puede referirse más ampliamente a las industrias cuya operación depende en gran medida de datos y análisis a gran escala. Prácticamente todos los negocios, excepto los más pequeños, eventualmente formarán parte de la “industria del Big Data”, del mismo modo que prácticamente todas las empresas usan computadoras y tecnología de comunicaciones en la actualidad.

Infraestructura de datos [Fuente: The Open Data Institute]

Consiste en activos de datos, en las organizaciones y sistemas que los operan y mantienen, y en procedimientos y guías que describen cómo usar y administrar los datos. Una infraestructura confiable de datos cuenta con financiamiento sostenible y supervisión que brinda orientación para maximizar el uso y el valor de los datos al satisfacer las necesidades de la sociedad. La infraestructura de datos incluye tecnología, procesos y organización.

Lago de datos [Fuente: PricewaterhouseCoopers]

Es un enorme repositorio que permite el almacenamiento a bajo costo de grandes cantidades y variedades de datos sin procesar y procesados, tanto estructurados como no estructurados. A diferencia del enfoque rígido y predefinido del modelo de datos relacionales, los esquemas de datos utilizados evolucionan continuamente en función de las necesidades de la organización. La flexibilidad de estos repositorios permite muchos usuarios dentro de un entorno dinámico. El término lago de datos generalmente se asocia con el almacenamiento de objetos orientado a Hadoop en el que los datos de una organización se cargan en la plataforma Hadoop y luego se aplican las herramientas de análisis de negocio y minería de datos donde estos son almacenados.

Sin embargo, los lagos de datos también se pueden usar de manera efectiva sin incorporar Hadoop en función de las necesidades y los objetivos de la organización. El término lago de datos se usa cada vez más para describir cualquier conjunto de datos de gran tamaño en el que los requisitos de esquema y datos no se definen hasta que se consultan los datos.

Privacidad y protección de datos [Fuente: Privacy International]

La privacidad es un derecho fundamental. Nos permite crear barreras y gestionar límites para protegernos de interferencias injustificadas en nuestras vidas, lo que nos permite negociar quiénes somos y cómo queremos interactuar con el mundo que nos rodea. La privacidad nos ayuda a establecer límites para decidir quién tiene acceso a nuestros lugares, cosas y espacio físico, así como nuestras comunicaciones y nuestra información. Más de 130 países tienen declaraciones constitucionales sobre la protección de la privacidad.

Un elemento importante del derecho a la privacidad es el derecho a la protección de los datos personales. La protección de datos se puede entender como un derecho, como leyes o como los mecanismos técnicos implementados para proteger este derecho. Al entenderse como ley, esta se diseña para proteger la información personal, que se recopila, procesa y almacena, por medios “automáticos” o que pretende formar parte de un sistema de archivo. En las sociedades modernas, para facultar a los ciudadanos a controlar su información y protegerlos de los abusos, es esencial que las leyes de protección de datos restrinjan y configuren las actividades de las empresas y los gobiernos. Más de 100 países ahora tienen alguna forma de ley de privacidad y protección de datos.

A. INTRODUCCIÓN

1. ALCANCE

Este es el segundo documento preparado y presentado como parte del apoyo brindado por MIT y Data-Pop Alliance para el diseño e implementación de la estrategia de Big Data de Colombia. Sigue y se basa en el Documento 1, que analizó las características principales de las estrategias gubernamentales existentes relacionadas con Big Data, mientras que aclaraba la conceptualización operativa de Big Data utilizada en todo este proyecto.

Específicamente, el Documento 1 formuló Big Data no sólo como datos masivos o grandes conjuntos de datos, sino como un fenómeno *tecnopolítico* o socio-tecnológico que emerge del surgimiento y desarrollo de un ecosistema compuesto por las 3 C de Big Data: Migajas, o las materias primas; Capacidades, la infraestructura que lo rodea y permite su explotación; y Comunidad (y Cultura) o la superestructura de actores e incentivos que rigen todo el sistema.

Como tal, el objetivo general de la consultoría y de una estrategia de Big Data no es simplemente, proporcionar pistas sobre formas de ‘analizar’ los nuevos grandes conjuntos de datos complejos resultantes de la transición de analógico a digital, sino de dar forma y aprovechar todo el ecosistema que se deriva de él, de forma que refleje y fomente los objetivos de desarrollo. Las amplias diferencias entre estos dos enfoques deberían ser evidentes por sí mismas, y se aclararán más en el Documento 4¹, pero para clarificar el objetivo y los propósitos de este documento basta decir que una estrategia sólida de Big Data puede implicar no usar Big Data en ciertas instancias.

Para el futuro de este ecosistema el sector privado en general y la industria de “Big Data” en particular, son fundamentales, ya que son los principales productores y recolectores de Big Data “crumbs (o migajas)”, aquellos donde la mayoría de las “capacidades” de Big Data se encuentran, y algunos de los miembros más poderosos e influyentes de la Comunidad Big Data. Los gobiernos de todo el mundo que se interesan en impulsar y dar forma a Big Data en la búsqueda de sus objetivos de desarrollo deberían centrar su atención en la incipiente industria Big Data.

¿Qué es la “industria de Big Data”?

La industria de Big Data es el conjunto de actores y actividades que generan ingresos a partir de datos, en uno o más puntos dentro de la cadena de valor de datos; incluye seguridad, servicios, infraestructura de almacenamiento, redes, infraestructura de centros de datos, herramientas de minería de datos, aplicaciones de big data, análisis, etc. También puede referirse más ampliamente a las industrias cuya operación depende en gran medida de datos y análisis a gran escala. Prácticamente todos los negocios, eventualmente formarán parte de la “industria del Big Data”, del mismo modo que prácticamente todas las empresas usan computadoras y tecnología de comunicaciones en la actualidad. Esto se debe a que los datos permiten un funcionamiento más eficiente y un mejor descubrimiento y explotación de oportunidades (por ejemplo, marketing). (Fuente: Alex Pentland)

La definición de Big Data presentada arriba implica la existencia de diferentes tipos de empresas:

1) las que toman sus decisiones primordialmente basadas en análisis de datos, yendo más allá del instinto y la experiencia de sus gerentes, 2) las que, por el contrario, no explotan sus datos y se basan en otros factores distintos de productividad para tomar decisiones, y 3) las empresas que proveen de servicios de tecnología de datos (en sus muchas formas) para que otras empresas u organizaciones aprovechen los beneficios de los datos.

Antes de pasar al resto de este documento, se deben aclarar algunos puntos conceptuales y definiciones adicionales para aclarar el objeto, alcance, objetivo(s) y estructura de este documento:

- **Tema:** este documento se refiere al sector o industria de Big Data como su principal objeto de análisis, para lo cual proporcionamos una definición de trabajo, señalando el desafío inherente en distinguir claramente qué empresa puede estar dentro o fuera del sector;
- **Alcance:** además, este documento se refiere a las diversas palancas y canales - desde incentivos fiscales a lineamientos regulatorios, marcos legales a intervenciones fiscales, hasta otros medios como abogacía - a través de los cuales un gobierno moderno como Colombia puede estimular este sector y su contribución a la economía del país. En otras palabras, su alcance es una política pública ampliamente entendida dirigida a fomentar una industria y crecimiento económico a través de ella;

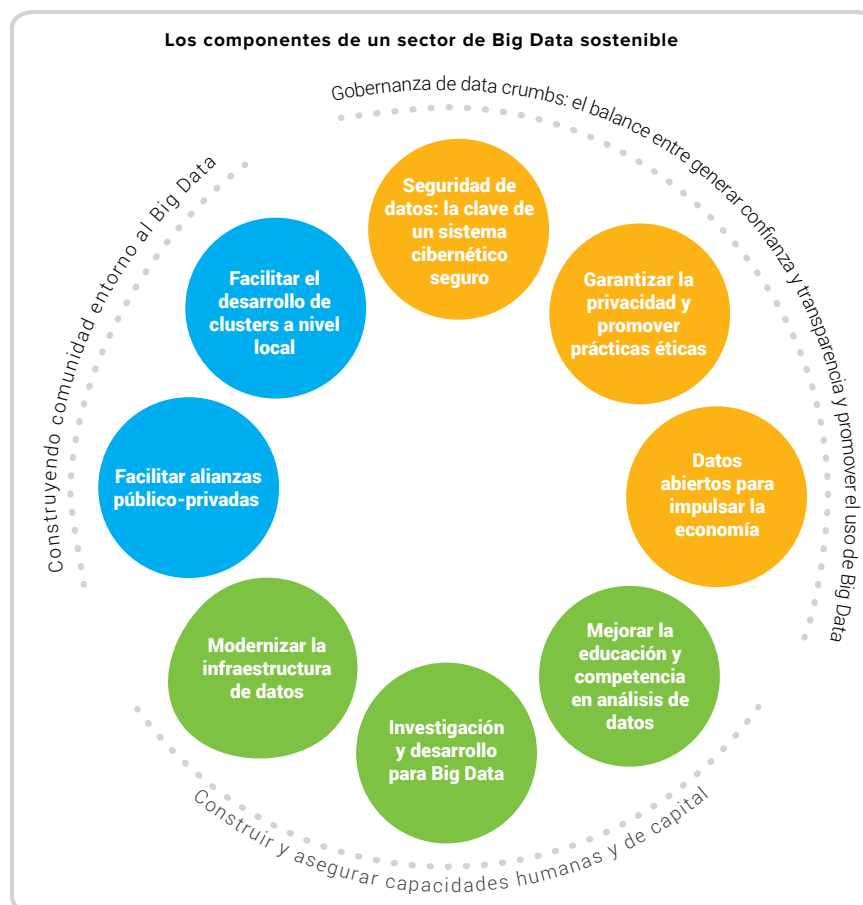
¹ Este documento contempla el desarrollo de la estrategia de Big Data y propone acciones concretas para su implementación.

² Estas empresas se pueden denominar también Data-Driven Companies o Datavores como se denominan en Reino Unido.

³ Estas empresas se pueden también denominar Dataphobes.

- **Objetivo:** este documento no tiene como objetivo hacer recomendaciones específicas de políticas públicas para promover la industria de Big Data colombiano, que es (uno de) los objetivos del Documento 4, sobre la base del análisis de la situación colombiana presentado en el Documento 3⁴; más bien, su objetivo es describir y analizar por qué, cuándo, cómo y, cuando sea posible, con qué resultados, los gobiernos de otros países seleccionados han utilizado algunas de estas palancas y canales, para fomentar la industria de Big Data. Esto, de la manera más organizada y legible posible;
- **Estructura:** la articulación del documento se hace a través de la lente analítica del marco 3C, específicamente organizando las iniciativas e intervenciones de política pública en términos de si (principalmente) apuntan a contribuir a la industria de Big Data y al crecimiento económico por (1) mejorar la gobernanza y el uso de Crumbs, así como otros tipos de datos, como datos abiertos, (2) mejorar el capital y las capacidades humanas y físicas, y (3) fomentar alianzas y clusters, respectivamente. Más precisamente, cada sección se divide en subsecciones, de la siguiente manera:
 - 1. Gobernanza de data crumbs: encontrar el equilibrio entre generar confianza y transparencia, y promover el uso de Big Data**
 - 1.1. Seguridad de datos: la clave de un sistema cibernético seguro
 - 1.2. Garantizar la privacidad y promover prácticas éticas
 - 1.3. Datos abiertos para impulsar la economía
 - 2. Construir y asegurar capacidades humanas y de capital**
 - 2.1. Mejorar la educación y las competencias en análisis de datos
 - 2.2. Investigación y desarrollo para Big Data
 - 2.3. Modernizar la infraestructura de datos
 - 3. Construyendo comunidad entorno al Big Data**
 - 3.1. Facilitar alianzas público-privadas
 - 3.2. Facilitar el desarrollo de clusters a nivel local

Figura 1: Los componentes de un sector de Big Data sostenible y las 3 Cs



Fuente: Elaboración propia

2. METODOLOGÍA

La metodología se desarrolló en dos partes: la primera fue establecer el marco analítico de la investigación y la segunda la revisión y clasificación de los principales casos o ejemplos de implementación de estrategias para estimular el naciente sector de Big Data, teniendo en cuenta la primera parte. El marco analítico que se utiliza es el que propone Data-Pop Alliance sobre los tres componentes del Big Data: Crumbs (migajas), Capacidades, y Comunidad y los asocia con áreas de acción puntuales.

En el caso de la primera C, Crumbs, se busca entender los mecanismos que llevan a una mejor gobernanza para el uso de Big Data, pasando por la creación de una sociedad donde hay confianza y transparencia. Se establecen como elementos centrales el fortalecimiento de la seguridad del espacio cibernético, la garantía de la privacidad y promoción del uso ético del Big Data, y el fomento de los datos abiertos.

Con respecto al componente de capacidades, el interés se encuentra en la manera de entender cómo se construye y asegura el capital humano necesario para manejar las necesidades recientes asociadas a la revolución de datos que actualmente se vive. En este componente, se tienen en cuenta líneas de acción asociadas a educación y competencias en análisis de datos, investigación y desarrollo, y modernización de la infraestructura de datos.

Finalmente, la última C, de comunidades, busca asociar los anteriores componentes en un marco integral que permita construir comunidad de manera sostenible. Las líneas de acción que alimentan este marco son las alianzas público privadas basadas en datos y el desarrollo de clusters locales que empoderen a dichas comunidades.

En cada uno de estos componentes, se desarrolló una metodología cuantitativa y cualitativa, que recoge las experiencias más relevantes de los países seleccionados. El criterio de selección se basó en parte en los resultados del documento 1 de esta consultoría: “Estado del arte y análisis comparativo de las estrategias de Big Data implementadas por los gobiernos”, en donde se tuvieron en cuenta indicadores que representan el grado de conectividad, avance tecnológico y capacidad para hacer negocios, estos índices fueron: Doing Business, Open Data Barometer, el Indicador de Capacidad Estadística, el Índice de Competitividad Global y el Índice de Disposición a la Conectividad (Network Readiness Index-NRI). Adicionalmente, se tuvo en cuenta el índice de participación del sector privado en proyectos públicos del Banco Mundial. A partir de ahí mapeamos las principales estrategias que buscan estimular el sector del Big Data.

B. ÁREAS DE ACCIÓN CLAVE

1. GOBERNANZA DE DATA CRUMBS: EL BALANCE ENTRE GENERAR CONFIANZA Y TRANSPARENCIA Y PROMOVER EL USO DE BIG DATA

El aumento de actividades económicas intensivas en datos (incluyendo “data crumbs” y datos abiertos) y la adopción creciente de tecnologías digitales (incluyendo computación en la nube, comercio en línea etc.), abren múltiples oportunidades pero se acompañan de un cambio en la escala y el alcance de una serie de riesgos, relacionados en particular con la protección de datos y la privacidad. Como veremos más adelante, estos riesgos tienen consecuencias en términos económicos y sociales.

Asimismo, las soluciones exploradas tienen a la vez que garantizar estándares de protección de datos y privacidad, y crear las condiciones para que prospere la innovación. En esta parte, exploramos un equilibrio entre protección y facilitar el crecimiento. Por un lado discutimos la promoción de medidas de protección de datos, teniendo en cuenta las preocupaciones de los usuarios individuales hasta las grandes organizaciones y la generación de confianza, a través del establecimiento de un marco legal que garantice la privacidad de los ciudadanos. Por otro lado, evaluamos la adopción de prácticas éticas que orienten el uso de Big Data por las organizaciones y por último medidas para impulsar la utilización de datos abiertos.

1.1. PROTECCIÓN DE DATOS: LA CLAVE DE UN SISTEMA CIBERNÉTICO SEGURO

La protección de datos es un componente central de un sistema cibernético seguro, que se define como un sistema que funciona como los usuarios esperan que funcione, sin acciones ni consecuencias no deseadas, y respetando los derechos del usuario sobre sus datos (incluyendo el respeto a la privacidad, la propiedad y el acceso a los datos). En el marco de la ‘industria de Big Data, crece la generación de “migajas de datos” (“data crumbs”), aumentando la cantidad de datos expuestos a infracciones de seguridad y filtración de datos personales (“data breaches”). Asimismo, el uso y la producción creciente de servicios digitales hacen de las amenazas y los ataques cibernéticos desafíos particularmente importantes para las organizaciones.

⁴ Este documento detalla el diagnóstico de la situación actual del Estado Colombiano en torno al Big Data.

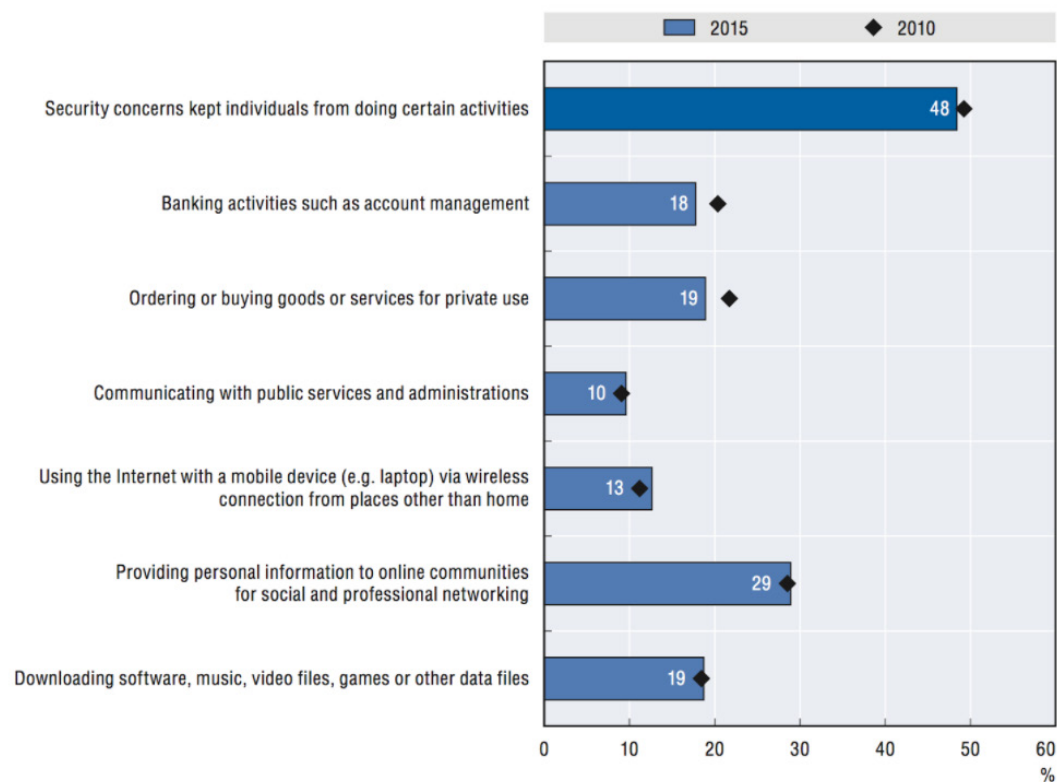
⁵ OECD (2017), OECD Digital Economy Outlook 2017, OECD Publishing, Paris. Disponible: <http://dx.doi.org/10.1787/9789264276284-en>

No existen estadísticas oficiales, fuentes de datos o metodologías reconocidas para medir las ocurrencias de ataques cibernéticos y sus costos asociados.⁵ Algunos estudios hacen estimados interesantes, pero tienen que ser tratados con cuidado. Por ejemplo, el estudio más reciente realizado por el Instituto Ponemon⁶ indica que en Estados Unidos el costo económico de los ataques cibernéticos ha crecido en los últimos 3 años, aumentando progresivamente de USD 11,51 en 2013 a USD 17,36 millones en 2016.

De acuerdo con la OCDE, “los riesgos de seguridad tienen efectos en la reputación, las finanzas y los activos físicos de las organizaciones, consecuentemente afectando su competitividad, habilidad para innovar y posición en el mercado”. El estudio realizado por el Instituto Ponemon citado arriba⁷, indica que la consecuencia más costosa para las organizaciones es la pérdida o robo de informaciones (en particular informaciones confidenciales y/o sensibles), seguida por las perturbaciones e interrupciones provocadas en las actividades, incluyendo la disminución de la productividad de los empleados y las fallas en los procesos operativos y comerciales. Los problemas de seguridad de datos afectan también el comportamiento de los consumidores frente a una marca o producto. El ‘grado’ de afectación puede variar dependiendo del contexto, es decir, si es por un distanciamiento de los consumidores, o si es por daños a la ‘imagen’ percibida de una empresa.⁸

Además de los costos para la organización misma, el análisis de la OCDE (2017) indica que los ataques a las empresas tienen efectos colaterales dentro de y más allá de su sector, y son mejor concebidos como desafíos intersectoriales. Los individuos también sufren las consecuencias de las filtraciones de datos, eventos que pueden generar ansiedad general y dudas sobre el concepto de compartir sus datos. La confianza de los ciudadanos y consumidores en las empresas se ve particularmente

Figura 2: Efectos de las preocupaciones de seguridad en las actividades de los individuos



Source: Eurostat, *Digital Economy and Society* (database), <http://ec.europa.eu/eurostat/web/digital-economy-and-society/data/comprehensive-database> (accessed March 2017).

Fuente: OECD (2017), *OECD Digital Economy Outlook 2017*, OECD Publishing, Paris.

⁶ Ponemon Institute. 2016 Cost of Cyber Crime Study & the Risk of Business Innovation, Benchmark Study of 237 Global Companies Ponemon Institute October 2016. Disponible: <https://software.microfocus.com/en-us/asset/2016-cost-cyber-crime-study-risk-business-innovation-ponemon-institute-cyber-security-analysis>

⁷ Ibid.

⁸ Acquisiti, Alessandro. “The Economics of Personal Data and the Economics of Privacy”. Carnegie Mellon University, Research Showcase @ CMU, (2010)

afectada por las brechas de seguridad. De acuerdo con un estudio de Eurostat (2017), casi la mitad de los ciudadanos han dejado de hacer ciertas actividades en línea por tener preocupaciones de seguridad (ver Figura 2), lo que afecta a su vez a las empresas. Además, otra encuesta del Instituto Ponemon⁹ calcula que los costos¹⁰ de las violaciones de datos ascendieron a \$6,65 millones de dólares por violación en 2008. En los últimos años esta cantidad ha continuado creciendo. Más allá de una disminución en la fe de un cliente o en la compra de una determinada compañía, se ha demostrado que tales violaciones de datos conducen a una disminución en la probabilidad de que un consumidor se involucre en ciertas actividades en general.

Además, los riesgos en términos de seguridad de datos y seguridad digital frenan la adopción de tecnologías digitales con alto potencial de rendimiento, como el ‘e-commerce’ (en particular e-commerce móvil) y la computación en la nube, que permanecen debajo de su potencial de acuerdo con la OCDE (2017). En Finlandia, Francia y Luxemburgo, más del 50% de las empresas no utilizan Internet móvil en todo su potencial debido a problemas de seguridad, aunque, como en el caso de Finlandia, más de un tercio de todas las oficinas necesitan una conexión móvil para su operaciones comerciales. En la Unión Europea, casi el 30% de las empresas no utilizan la nube, debido ante todo a preocupaciones sobre los riesgos de filtración de datos (personales), y secundariamente a incertidumbres sobre la ubicación de los datos (OCDE, 2017).

Siendo así, para promover el uso de Big Data con el objetivo de generar un mayor impacto en la industria y en la economía del país, es clave el refuerzo de la seguridad de datos y seguridad digital. El Estado no puede responder a los riesgos y garantizar un espacio cibernético seguro por sí sólo pero puede facilitar su promoción. Ejemplos de medidas y prácticas al alcance del gobierno son exploradas adelante.

Crear herramientas y métodos para comprender el riesgo sistémico de ciberseguridad

Como mencionamos al principio de esta parte, todavía no existen estadísticas oficiales, fuentes de datos o metodologías reconocidas para medir las ocurrencias de ataques cibernéticos y sus costos asociados. Una de las razones es que las organizaciones no están inclinadas a compartir información potencialmente perjudicial, y en muchos casos no suelen reportar los incidentes, especialmente cuando no están legalmente constreñidas a hacerlo (OCDE, 2017). La medición de las ocurrencias y costos de estos ataques permitirían a mínima orientar la toma de decisiones y una mejor gestión del riesgo de ciberseguridad. El estado puede jugar un rol para promover la generación de datos y estadísticas sobre el tema, aprovechando por ejemplo proyectos que se están desarrollando en la academia para impulsar la provisión de estos datos por las empresas. Por ejemplo, este año, el interdisciplinario Internet Policy Research Initiative (IPRI) del MIT invirtió USD 1.5 millones para apoyar proyectos de investigación relacionados con la ciberseguridad, a fin de desarrollar recomendaciones para los legisladores en tomar decisiones informadas y asegurar que los incentivos apropiados estén en su lugar para reducir la amenaza de ataques catastróficos. Entre ellos, el proyecto de Andrew Lo (Sloan) & Vinod Vaikuntanathan (CSAIL) propone desarrollar herramientas y métodos para comprender el riesgo sistémico de ciberseguridad, con el fin de proporcionar mejor información y entendimiento de estas dinámicas. El objetivo del proyecto es crear una “plataforma segura de computación multiparte que otorgue a las empresas la capacidad de agrupar datos encriptados preservando la confidencialidad, y permita mapear los vínculos entre las empresas y calcular estadísticas resumidas”.¹¹

¿Cuáles son los principios básicos para garantizar el funcionamiento de un espacio cibernético seguro?

De acuerdo con Alex Pentland, existen tres principios básicos de la seguridad cibernética: (1) saber “quién, qué, cuándo y dónde”, (2) tener una estrategia de defensa basada en la tecnología y diversificación del riesgo, y (3) promover prácticas seguras entre todos los actores, incluyendo los ciudadanos y promover la utilización de código y sistemas abiertos.

El principio básico de seguridad cibernética es saber siempre “**quién, qué, cuándo y dónde**”, y tener la posibilidad de retroceder y reconstruir eventos. En consecuencia, la seguridad en el ciberespacio requiere de (i) una fuerte identidad digital, (ii) autenticación y permisos confiables, y (iii) registro y auditoría incorruptibles. En este sentido, promover la utilización de tecnologías como la verificación de la identidad de las personas y empresas que hacen transacciones en línea, es una de las herramientas que los gobiernos pueden implementar. Algunos países, en especial en Europa están haciendo progresos en ese sentido. Reino Unido ha adoptado un sistema que permite verificar la identidad de los usuarios contra los detalles que se encuentran en una serie de fuentes, incluyendo agencias de referencia crediticia, facturas de servicios públicos, permisos de conducción y facturas de telefonía celular. El sistema no retiene esas piezas de información, y las compañías de verificación

⁹ Ponemon, D. (2009). Fourth Annual US Cost of Data Breach Study; Benchmark Study of Companies. Traverse City, MI: Ponemon Institute.

¹⁰ Sumando el valor de ventas, los costos legales, las reparaciones a los consumidores, etc.

¹¹ Gordon, Rachel. “New Funding Enables Work On Internet Policy And Cybersecurity For Key Infrastructure”, MIT Computer Science and Artificial Intelligence Laboratory News. May 19, 2017. http://www.csail.mit.edu/new_funding_enables_work_on_internet_policy_and_cybersecurity_for_key_infrastructure

no conocen los servicios que están siendo usados. Solo un número de teléfono fijo o móvil es registrado con el propósito de enviar códigos de verificación solicitados en los pasos siguientes. Cuando las personas ingresan¹², deben proveer una identificación y contraseña, y verificar su identidad ingresando un código de ingreso enviado vía telefónica.

El segundo principio básico es **tener mecanismos de defensa profundos y variados (nunca confíes en un mecanismo, nunca pongas todo en un solo lugar)**. En este sentido, se recomienda que las organizaciones introduzcan nuevas prácticas de gestión de riesgos (en particular para las Pymes). En el Informe anual de amenazas de seguridad¹³, publicado por la empresa Dell en 2015 y basado en datos de más que un millón de sensores en 2000 países, se explica que los ciberdelincuentes están constantemente innovando y diversificando sus técnicas y herramientas. El informe describe que los métodos de ataque evolucionan constantemente, y que se desarrollan cada mes nuevos tipos de malware que pueden atacar los sistemas de pago de todo el mundo.¹⁴

Como resultado, las técnicas de seguridad han evolucionado para enfrentar el desafío, incluyendo la utilización de métodos de aprendizaje de máquinas (“machine learning”) para detectar problemas de manera más inmediata (a medida que ocurren) y eficiente, con el fin de mitigarlos rápidamente para minimizar el daño potencial. El laboratorio de Informática e Inteligencia Artificial del MIT ha desarrollado una plataforma llamada AI², que puede revisar miles de líneas de código necesarias para encontrar una potencial brecha de seguridad. Los desarrolladores de la plataforma enfatizan que no se puede depender completamente de una máquina para verificar un ataque, un operador humano ayuda y verifica su trabajo; pero la gran potencia informática del sistema puede localizar rápidamente posibles amenazas de seguridad en una medida que ningún humano por sí solo puede.¹⁵ Mediante el uso de plataformas como esta, las empresas privadas y los gobiernos por igual pueden responder a la naturaleza en constante cambio de los ciberataques.

El tercer principio básico es que las personas constituyen la mitad de todos los sistemas cibernéticos clave: **el 70% de los ataques tienen éxito debido a un error humano**. La mayoría de las fallas son descubiertas por utilizadores del sistema en vez de creadores del sistema. En consecuencia, es fundamental utilizar sistemas de código abierto y sistemas fácilmente auditables, para que muchos ojos puedan ayudar a encontrar problemas.

En un esfuerzo por reducir la incidencia de tales errores humanos, ha surgido una gran cantidad de empresas privadas de capacitación en seguridad digital. Ya existe una amplia gama de programas, dirigidos a grupos que van desde los trabajadores a nivel principiante hasta los directores ejecutivos, con sesiones en línea o en persona, e incluso aquellos creados por escritores de comedias para hacer que el material sea más agradable para los estudiantes.¹⁶ Algunas instituciones y compañías privadas también han desarrollado sus propios programas internos de capacitación. Entre ellos, la Universidad de Berkeley en California (UB) lanzó una guía de ‘mejores prácticas’ para sus estudiantes y comunidad. Su guía enfatiza la noción de responsabilidad colectiva, afirmando que “ningún individuo, empresa o entidad gubernamental es el único responsable de proteger el internet. Todos tenemos un rol en asegurar nuestra parte del ciberespacio, incluidos los dispositivos y las redes que utilizamos. Las acciones individuales tienen un impacto colectivo y cuando utilizamos el internet de forma segura, lo hacemos más seguro para todos.”¹⁷

Sin embargo, un desafío importante para una oficina gubernamental es la necesidad de difundir esta conciencia y educación entre todos los usuarios. Como la Comisión Federal de Comercio (FTC) de EU afirma en su guía sobre seguridad cibernética: “la seguridad sólida no es accidental.”¹⁸ Las instrucciones de 10 pasos describen en detalle cómo una empresa privada puede formular una estrategia segura de protección digital desde cero, incluso a los protocolos de administración de contraseñas, cómo monitorear redes internas, y cómo abordar vulnerabilidades que puedan surgir. Al ofrecer este asesoramiento de manera gratuita, permite que empresas de todos los niveles de desarrollo puedan accederlo. Como se ha descrito anteriormente, todos los sectores de la economía y, como consecuencia, de la sociedad, se ven afectados por la ciberseguridad y, por lo tanto, necesitan la misma protección; un organismo gubernamental es la única entidad que podría proporcionar ese consejo.

En esta parte vimos que los ataques cibernéticos tienen costos económicos y sociales. La promoción de un sistema cibernético seguro empieza con un mejor entendimiento del riesgo sistemático de ciberseguridad, lo que puede ser facilitado por un gobierno. Además, exploramos tres aspectos de la ciberseguridad en los cuales el gobierno tiene un rol a jugar: (1) saber “quién, qué, cuándo y dónde”, (2) tener una estrategia de defensa basada en la tecnología y diversificación del riesgo, y (3) promover prácticas seguras entre todos los actores, incluyendo a nivel individual - así como la utilización de código y sistemas abiertos.

¹² Para ingresar a este sistema, se debe tener más de 19 años, vivir en el Reino Unido y ser residente por más de 12 meses.

¹³ Dell Inc., “Dell Security Annual Threat Report”, 2015. Disponible: <http://www.enhacker.com/wp-content/uploads/2015/06/2015-dell-security-annual-threat-report-white-paper-15657.pdf>

¹⁴ Ibid.

¹⁵ Conner-Simons, Adam. “System predicts 85 percent of cyber-attacks using input from human experts”, MIT News. April 18, 2016.

¹⁶ Para ejemplos y más detalles de los programas de entrenamiento ver a Digital Defense Inc, Inspired eLearning, Wombat Security o Know-Be4.

1.2. GARANTIZAR LA PRIVACIDAD Y PROMOVER PRÁCTICAS ÉTICAS

La privacidad es un derecho humano, parte de la Declaración Universal de Derechos Humanos de 1948 y ha sido reconocido como derecho fundamental en la mayoría de países del mundo. Los datos personales, fuertemente asociados a la privacidad, pero sujetos de una protección autónoma, componen una parte importante de las ‘migajas’ de Big Data (registros de datos de llamada (CDR), huellas en línea (e.j. cookies del navegador, direcciones IP, historias de búsqueda), redes sociales, transporte (aplicaciones para transporte compartida), finanzas (mercados en línea, transacciones con tarjeta de crédito, etc). Esto, plantea una serie de desafíos reales asociados a la privacidad y a la protección de datos, entre ellos: riesgos de vigilancia (e.j. la posibilidad de gobiernos o empresas de rastrear continuamente individuos) y discriminación (e.j. discriminación basada en datos médicos), la utilización de datos personales para fines lucrativos sin consentimiento informado y la falta de transparencia y (posibles) sesgos de los algoritmos (ICO, 2017).¹⁹ Por otro lado, el uso de métodos de análisis Big Data permite inferir información sensible a partir de datos que pueden parecer triviales al principio, como el comportamiento de compra individual o el consumo de electricidad; esto es conocido como “profiling”. Potencialmente, tal información puede ser utilizada para clasificar injustamente a una persona y hacer que sea discriminada con base en rastros de datos.

Actualmente, más de 100 países poseen una ley sobre privacidad de datos, incluyendo Colombia. La Oficina del Comisionado de la Información (ICO) de Reino Unido define la privacidad como un “derecho habilitante” (“an enabling right”), es decir no solamente un fin en sí mismo, sino un medio de asegurar que Big Data cumpla con funciones benéficas para la sociedad, a través de la generación de confianza y transparencia entre los actores del ecosistema de Big Data, en particular entre los sujetos de datos (ciudadanos en particular) y los utilizadores de datos (empresas, gobiernos).²⁰ Actualmente, “el mercado no provee suficiente protección a la privacidad porque los consumidores no creen que tienen el control de su privacidad o que las compañías realmente pueden protegerlos.”²¹ Más allá de atender las preocupaciones de los ciudadanos, la transparencia y el acceso a los datos personales producen beneficios en productividad y eficacia de negocios y programas públicos. De acuerdo con Privacy International,²² la única forma en que los ciudadanos y los consumidores pueden tener confianza en el gobierno y las empresas es a través de sólidas prácticas de protección de datos, con una legislación eficaz para ayudar a minimizar el monitoreo innecesario por parte de la burocracia y regular la vigilancia de las empresas”. Analizar los derechos de propiedad de los datos es el primer paso que cualquier gobierno debe iniciar a la hora de establecer una regulación robusta al respecto. Este análisis necesariamente debe partir de la pregunta ¿a quién le pertenecen los datos?, y en consecuencia, ¿quién puede utilizar estos datos y bajo cuáles condiciones? Un gobierno debe además tener claro cuál es el mapa actual de su regulación, sus actores y sus dinámicas, para luego identificar con claridad en dónde están los vacíos y por último, priorizar acciones.

En esta parte, argumentamos que el rol del Estado es facilitar la generación de confianza y transparencia, garantizando la privacidad de los ciudadanos, y al mismo tiempo incentivar la industria promoviendo prácticas éticas para orientar (e impulsar) el uso de Big Data por las empresas.

1.2.1 GARANTIZAR LA PRIVACIDAD DE LOS CIUDADANOS: EL EJEMPLO DEL GDPR

Existen ya en el mundo diferentes aproximaciones sobre la regulación apropiada para asegurar la protección de datos sin disminuir los esfuerzos en materia de competitividad e innovación. A continuación analizamos los aportes del Reglamento General de Protección de Datos de la Unión Europea (GDPR).²³ El GDPR, adoptado en 2016 por los Estados miembros de la Unión Europea (EU) que entrará en vigor en mayo de 2018, sustituyendo la Directiva de Protección de Datos (DPD) de 1995, es considerado uno de los conjuntos de normas más completos que se está adoptando sobre la protección de datos en la era digital. Su principio fundamental es “devolver a los ciudadanos el control sobre sus datos personales”²⁴ y debe ser aplicada a cualquier compañía, independientemente de su origen, que detenga datos sobre ciudadanos residentes en Europa.

Asimismo, el GDPR refuerza los derechos de los ciudadanos, con la creación del derecho a la portabilidad de los datos (particularmente interesante desde el punto de vista del consumidor, ya que implica poder cambiar fácilmente de servicio, probar servicios equivalentes, etc.), con el hecho de que el derecho a ser olvidado se vuelve más explícito, con condiciones estrictas para obtener el consentimiento para el procesamiento de datos personales. Introduce también nuevas disposiciones para todo lo que sea ‘profiling’ y decisiones automatizadas (si producen “efectos legales” o “significativos”) y la obligación de explicar hasta cierto punto la “lógica” utilizada para tomar ciertas decisiones automatizadas.

Para las empresas, la principal promesa es la armonización de las leyes en los 28 estados miembros. No obstante, el GDPR deja a los Estados Miembros un margen de maniobra y parece estar avanzando hacia diferencias nacionales significativas. La implementación de “la ventanilla única” y el “mecanismo de conveniencia” (que prometía la simplificación de los

¹⁷ Conrad, Larry y Rivers, Paaul. “Protect Yourself: Staying Safe in Cyberspace”, Technology @ Berkeley. October 6, 2015. <https://technology.berkeley.edu/news/protect-yourself-staying-safe-cyberspace>

¹⁸ Federal Trade Commission, “Start with Security: A Guide for Business”. Disponible: www.ftc.gov/tips-advice/business-center/guidance/start-security-guide-business

¹⁹ Information Commissioner’s Office (ICO). “Big data, artificial intelligence, machine learning and data protection”, Data Protection Act and General Data Protection Regulation. Versión 2.2, 2017.

²⁰ Ibid.

²¹ Banco de Desarrollo de América Latina, 2013, Technology OECD Directorate for Science and Innovation, “Data-driven innovation for growth

procedimientos de ejecución para las multinacionales) es, por el momento, muy incierta. El GDPR insiste en gran medida en las nociones de rendición de cuentas de las empresas, introduciendo la entrada en vigor de sanciones. En particular multas, basadas en un mecanismo escalonado por incumplimiento, son impuestas por la “Autoridad Supervisora” centralizada (DPA)²⁵ dependiendo de la naturaleza, la gravedad y la duración de la infracción. Hay multas graves que pueden implicar hasta el 4 por ciento del volumen de negocios global anual de una empresa y 20 millones de euros por incumplimiento a principios como el de consentimiento. Otras sanciones contemplan un 2 por ciento del volumen de negocios global anual y 20 millones de euros. El establecimiento de los montos de estas multas hace que las empresas estén buscando desde ya prepararse para la implementación del GDPR.

La mayoría de Estados de la Unión Europea están realizando acciones para ayudar a las empresas a prepararse para enfrentar los cambios que deberán adoptar a partir de mayo de 2018 (cuando la GDPR entrará en vigor) en materia de protección de datos personales. En Reino Unido, por ejemplo la Oficina del Comisionado de la Información (ICO), como cuerpo regulador independiente en ese país, es responsable de su implementación y por tanto ha establecido una guía, blogs y eventos para asegurarse que las empresas han comprendido las implicaciones de la nueva Regulación. Se estima que la implementación del GDPR, en particular la aplicación de dispositivos informativos y de acompañamiento dirigidos directamente a las empresas, está asegurando desde ya cambios relevantes en la manera como se utilizan y explotan los datos.

El impacto que tendrá el GDPR sobre el uso de Big Data, incluyendo tecnologías de inteligencia artificial y machine learning, aún es incierto. La preparación de la nueva política para abordar las complejidades de la industria de Big Data se evaluó en un estudio²⁶ que indica que la implementación del GDPR se tornaría más difícil e ineficiente el proceso de análisis de Big Data. Por un lado, expertos han criticado el GDPR al señalar que su cumplimiento por parte de las empresas, va a implicar un incremento en el número de trámites y papeleo, así como mayores esfuerzos en el almacenamiento y registro de datos, por ejemplo al imponer “obligaciones onerosas de rendición de cuentas”²⁷.

Por otro lado, en las conferencias académicas y de privacidad la discusión todavía no está muy avanzada. Las disposiciones sobre la “toma de decisiones automatizada” (Artículo 22) pueden tener consecuencias sobre el uso de algoritmos, y asimismo la necesidad de consentimiento aplicada a las decisiones automatizadas (considerada necesaria para las decisiones que producen “efectos legales” o que tienen “efectos similarmente significativos” en el individuo). El impacto del GDPR dependerá de la interpretación de las autoridades de protección de datos al respecto. La guía emitida el mes pasado por el Grupo de Trabajo del Artículo 29 (documento actualmente abierto para consulta), adopta una definición amplia del término “efectos similarmente significativos”.

La necesidad de consentimiento válido para actividades de ‘profiling’, definidas como “cualquier forma de procesamiento automatizado de datos personales que consiste en el uso de datos personales para evaluar ciertos aspectos personales relacionados con una persona física, en particular para analizar o predecir aspectos relacionados con el desempeño de esa persona física en el trabajo, situación económica, salud, preferencias personales, intereses, fiabilidad, comportamiento, ubicación o movimientos” (Art 4(4)), también puede afectar la utilización de métodos analíticos Big Data.

Los trabajos de Sandra Wachter, Luciano Floridi y Brent Mittelstadt, tratan de determinar cómo el GDPR impone una forma de transparencia en los algoritmos o un derecho a la explicación. En su artículo (2016), los autores discuten que aunque la legislación venidera (que entrará en vigor en 2018) propone garantizar dicha transparencia, en realidad su viabilidad es restringida por el alcance y la falta de especificidad de la ley. Como se enuncian en el artículo, “el GDPR solo obliga que los sujetos de los datos reciban significativa, pero adecuadamente limitada, información... y plantea dudas sobre la protección que en realidad se ofrece a los sujetos de datos.”²⁸ La implementación de transparencia verdadera requerirá protección más allá de lo que el GDPR puede ofrecer.

Dado que el GDPR está en fase inicial de implementación, aún es muy a priori calificar y cuantificar exactamente el efecto que tendrá sobre el uso de Big Data. Sin embargo, es considerado uno de los conjuntos de normas más completos que se está adoptando sobre la protección de datos en la era digital. Asimismo, contiene aspectos relevantes que pueden servir de referencia como condiciones estrictas para obtener el consentimiento, la creación del derecho a la portabilidad de los datos, el refuerzo del derecho a ser olvidado y un fuerte sistema de rendición de cuentas de las empresas, introduciendo la entrada en vigor de multas. Sin embargo, todavía es incierto el efecto que algunos de estos aspectos tendrán en la práctica, por ejemplo las condiciones estrictas de consentimiento sobre el uso de algoritmos, por ejemplo para toma de decisiones automatizadas y actividades de ‘profiling’.

and well-being,” OECD Interim Synthesis Report, 2014.

²² Privacy International, “What is Data Protection?”. Disponible: www.privacyinternational.org/node/44

²³ European Union, “Regulation (EC) No 679/2016 of the European Parliament and of the Council of the 27 April 2016 The General Data Protection Regulation,” OJ L 119 (2016): 33–34.

²⁴ El GDPR señala los siguientes derechos para las personas: (1) el derecho a estar informado, (2) el derecho al acceso, (3) el derecho a la rectificación, (4) el derecho a enmendar, (5) el derecho a restringir procesos (que incluye “el derecho al olvido”, esto es, solicitar que sus datos sean borrados), (6) el derecho a la portabilidad de la información, (7) el derecho a objetar (incluyendo el derecho a objetar el marketing directo), (8) derechos en relación con tomas de decisiones y esbozos automatizados. En cuanto al consentimiento, las empresas deben asegurar que éste sea tan fácil de retirar como de otorgar y no puede ser inferido mediante silencio, casillas premarcadas o inactividad.

1.2.2 PROMOVER PRÁCTICAS ÉTICAS PARA ORIENTAR EL USO DE BIG DATA POR EL SECTOR PRIVADO

Mencionamos arriba los desafíos en términos de respeto a la protección de datos y respeto a la privacidad que acompañan el uso de Big Data. Al mismo tiempo, también se han expresado preocupaciones éticas sobre el uso de Big Data, incluyendo en términos de representatividad y justicia. Por ejemplo, en “Weapons of Math Destruction”, Cathy O’Neil denuncia los sesgos y la opacidad de algoritmos que pretenden cuantificar el riesgo de reincidencia y la solvencia crediticia, que terminan reforzando prejuicios y a menudo la desigualdad.²⁵ Para lograr un balance apropiado entre los beneficios y riesgos, se han desarrollado guías éticas para el uso de nuevas fuentes de datos.

Por ejemplo, en mayo de 2016, Reino Unido estableció un Consejo de Ética en Ciencias de Datos (Council of Data Science Ethics), en respuesta a una solicitud del Comité de Ciencia y Tecnología, con el objetivo de guiar a las personas en la realización de proyectos de ciencia de datos y generar confianza para innovar con Big Data. Al mismo tiempo, el Gabinete Ministerial publicó un “Marco Ético de Ciencia de Datos” destinado a proporcionar a los funcionarios públicos orientación sobre la realización de proyectos de ciencia de datos y la confianza para innovar con datos.

Este marco fija seis principios éticos para el desarrollo de iniciativas en datos: (i) empezar con un análisis de las necesidades y del beneficio público, (ii) usar los datos y herramientas que tengan un mínimo nivel de intrusión, (iii) crear modelos robustos de ciencias en datos, (iv) estar alerta de percepciones públicas, (v) contar con la mayor apertura y rendición de cuentas que sea posible, (vi) mantener los datos lo más seguro posible. Si bien el Comité de Ciencia y Tecnología había recomendado que este Comité introdujera sanciones penales por graves infracciones a la privacidad y que se realizarán auditorías obligatorias de la protección de datos de las autoridades locales; por lo menos en esta fase inicial, el nuevo Comité de Ética no lo hará y, en su lugar, concentrará sus esfuerzos en la revisión de las sanciones vigentes en el marco de la nueva normatividad de la Unión Europea sobre protección de datos y se centrará en auditorías realizadas de forma voluntaria.

Si bien el Comité Ético está recién establecido y aún es muy difícil analizar su pertinencia, las recomendaciones emitidas por el Comité de Ciencia y Tecnología incluyen realizar auditorías obligatorias a nivel local y emitir sanciones penales a infractores, aspectos aplicables al sector privado.

De cualquier manera, los órganos institucionales a establecer por los países para lidiar con los aspectos éticos de explotación de datos, denominándolos Comités Éticos o de otro modo, deben también considerar la participación de un conjunto de expertos de diversa índole que aporte su conocimiento en los casos donde la privacidad de los datos sea cuestionada. Estos expertos pueden contar con la legitimidad, desde su especialidad, para asegurar la protección de los datos personales. Para que este consejo sea establecido y legítimo, todos los actores clave de la sociedad deben estar representados. La academia puede liderar estos procesos junto con la sociedad civil, pero el gobierno debe también estar vinculado, en el sentido de escuchar y poner en práctica las recomendaciones que estos expertos establezcan, haciendo estos marcos éticos realmente efectivos.

En esta parte, “Garantizar la privacidad y promover prácticas éticas”, argumentamos que el rol del Estado es facilitar la generación de confianza y transparencia, garantizando la privacidad de los ciudadanos a través del establecimiento de un marco legal reforzado para atender a los desafíos, en términos de protección de datos y privacidad en la era digital. Al mismo tiempo, el gobierno debe explorar medidas que sigan impulsando la industria de Big Data. Entre ellas, exploramos la promoción de prácticas éticas para orientar (e impulsar) el uso de Big Data por las empresas, a través la creación de marcos éticos y órganos de implementación y consejos correspondientes.

1.3. DATOS ABIERTOS PARA IMPULSAR LA ECONOMÍA

La apertura de datos es una oportunidad para estimular el crecimiento económico y fomentar el ecosistema de Big Data. El Banco Mundial sostiene que los datos abiertos tienen un potencial económico significativo: en el 2014 se rastrearon nuevas compañías en Estados Unidos que usaban datos abiertos y se valoraron en 1 billón de dólares (para efectos de escala, corresponde aproximadamente al PIB de Guinea Bissau).

Como poseedor y publicador de datos abiertos, el gobierno está a la vanguardia de la apertura y promoción de éstos. En esta parte, nos enfocamos en tres medidas principales que puede tomar el gobierno para promover la utilización de datos abiertos en el ámbito de generar mayor impacto económico: (1) promover una mayor apertura de los datos públicos, a diferentes niveles del gobierno; (2) promover o regular la apertura de datos en otros sectores (en particular empresas y academia) y (3) impulsar el “emprendimiento de datos”, a través del apoyo a la innovación y la incubación de nuevas empresas.

²⁵ El GDPR prevé la creación de una “Autoridad Supervisora” centralizada denominada DPA. Esta figura permite a una empresa que está establecida en varios países tratar con una sola autoridad nacional de protección de datos. Esta figura también está relacionada al concepto de ventanilla única que pretende implementar el GDPR.

²⁶ Zarsky, Tal, Incompatible: The GDPR in the Age of Big Data (August 8, 2017). Seton Hall Law Review, Vol. 47, No. 4(2), 2017. Disponible: <https://ssrn.com/abstract=3022646>

²⁷ Overy, Allen. The EU General Data Protection Regulation. 2017.

²⁸ Wachter, Mittelstadt, Floridi. “Why a right to explanation of automated decision-making does not exist in the General Data Protection

Cómo se genera valor económico a través de los datos abiertos?

El reporte “How Government Can Promote Open Data And Help Unleash Over \$3 Trillion In Economic Value” de McKinsey, estima que los datos abiertos pueden ayudar a generar más de \$ 3 billones de dólares anuales para la economía global. Los datos abiertos ya están contribuyendo al crecimiento económico de varios países en el mundo, también están apoyando la creación y el fortalecimiento de nuevos mercados, empresas y empleos. El gobierno juega un papel relevante en la creación de valor a partir de datos abiertos, no sólo en su etapa de publicación.

De acuerdo con este mismo reporte McKinsey, se puede crear valor con los datos abiertos en las diversas empresas e industrias a través de tres formas: 1) para que las empresas tradicionales o nuevas startups no tecnológicas tomen decisiones, de la misma forma que los propios gobiernos usan datos abiertos para mejorar la toma de decisiones, lo mismo puede suceder para el vasto sector privado; 2) para generar nuevos productos o servicios que generan valor para los clientes de las empresas; y 3) para rendir cuentas en un mercado donde los consumidores requieren de más información y premian a las empresas transparentes: al liberar datos, las empresas pueden garantizar que sus acciones sean transparentes

1.3.1 PROMOVER UNA MAYOR APERTURA DE LOS DATOS PÚBLICOS, A DIFERENTES NIVELES DEL GOBIERNO

Uno de los mecanismos a disposición del gobierno para fomentar el uso de datos, en particular datos abiertos, por el sector privado, es fortalecer la oferta éstos, en términos cuantitativos y cualitativos.

En Corea del Sur, el gobierno ha impulsado una serie de medidas para promover los datos abiertos, permitiendo el desarrollo de muchas aplicaciones digitales a partir de los datos abiertos públicos. Uno de los sitios de Open Data más activos es el Seúl Open Data Plaza (data.seoul.go.kr) gestionado por el gobierno metropolitano de Seúl. En 2012, Seúl inició una iniciativa de datos abiertos compartiendo información pública con el objetivo de crear diversas oportunidades de negocios para el sector privado y desarrollar industrias de TI. Este portal es un canal en línea para compartir y proporcionar a los ciudadanos todos los datos públicos de Seúl, tales como horarios de autobuses en tiempo real, horarios de metro, lugares de servicios públicos Wi-Fi, instalaciones para personas discapacitadas, entre otros. La información registrada en Seúl Open Data Plaza se proporciona en el formato abierto de la API, y está diseñado para permitir que las personas, y especialmente las empresas puedan utilizarlo en la creación de nuevos negocios. Del mismo modo, las aplicaciones innovadoras en los sectores del transporte y de la salud muestran cómo el país está aprovechando el poder de los grandes datos.

Desde la Ley de Provisión y Uso Activo de Datos Públicos promulgada el 27 de junio de 2013 y aplicada desde el 31 de octubre de 2013, se han hecho más progresos. Esta Ley establece explícitamente el derecho de los ciudadanos a utilizar datos públicos abiertos y establece la obligación a las agencias del sector público a proporcionar datos públicos. Uno de sus objetivos principales es crear las bases para incentivar el uso de datos públicos por el sector privado.

Asimismo, la ley se acompaña de planes de implementación que abordan tres dimensiones principales: (1) evaluar el rendimiento de la provisión y uso de datos abiertos en el año anterior, (2) detallar el plan maestro del año, y (3) hacer un plan de gestión del presupuesto para el año. Desde la promulgación de la Ley en 2013, el Ministerio del Interior (MOI) y el Centro Nacional de Datos Abiertos (NIA) han buscado asegurar el cumplimiento de estos planes por las agencias nacionales y gobiernos locales, facilitando el entendimiento de la visión, los objetivos y las prioridades establecidas, así como proporcionando apoyo técnico. El cumplimiento de estos planes ha resultado en la apertura de 9,259 conjuntos de datos gubernamentales entre 2013 y 2015, es decir un incremento de la tasa de apertura de los datos públicos del 16.1% en 2013 a aproximadamente el 50% 2015.

1.3.2 PROMOVER O REGULAR LA APERTURA DE DATOS EN OTROS SECTORES (GOBIERNOS LOCALES, ACADEMIA, O EMPRESAS)

Los gobiernos junto con algunos organismos internacionales y asociaciones civiles han estado a la vanguardia de la apertura y proliferación de datos abiertos. No obstante, existen sectores clave para la sociedad que no han tenido el mismo compromiso con la apertura de sus datos.

Desde el 2014, el Govlab señaló a la apertura y accesibilidad de datos en el sector privado como la próxima frontera del Open Data. Conforme los gobiernos han adoptado la agenda de datos abiertos, los ciudadanos, y por tanto, los consumidores reclaman apertura en otros sectores como la empresa, la academia, y las organizaciones gubernamentales. Como se ha señalado en este documento, los datos se han convertido en la moneda de la economía moderna. Un estudio reciente (Noviembre 2017) publicado por el “Future of Privacy Forum”³⁰ proyecta que el volumen global de datos crecerá de aproximadamente 0,8 zettabytes (ZB) en 2009 a más de 35 ZB en 2020, la mayor parte generada en los últimos dos años y administrada por el sector corporativo. El rol del gobierno en la promoción de la apertura de datos corporativa es fundamental: por un lado, muchos de estos datos son cruciales para la seguridad de las personas (datos automovilísticos), y su salud (datos de la industria alimenticia).

Asimismo, el gobierno puede jugar un rol para incentivar las empresas a compartir sus datos de manera segura y respetuosa de la privacidad de los consumidores y ciudadanos. Por ejemplo, promover plataformas seguras para el intercambio de datos del sector privado, como lo hace el proyecto OPAL³¹ (se analiza en mayor detalle en la sección 3.1). OPAL proporciona los medios necesarios para convertir los “datos masivos” en “datos abiertos”: “abriendo sin exponerlos” datos recabados y almacenados por empresas privadas. Esto se logra proporcionando sólo indicadores agregados relevantes calculados sobre la base de “migajas” de datos a través de algoritmos abiertos que se ejecutan en los servidores y detrás de los firewalls de las compañías privadas, de una manera que se preserva la privacidad. Más allá de esta importante función, OPAL está diseñado para reflejar y fomentar las normas y estándares para la industria del Big Data y las economías y sociedades infundidas de datos de manera más amplia, como se discute en la tercera parte.

Las universidades y la academia (en sus diversas instituciones como las agencias de ciencia y tecnología) también tienen que dar un paso hacia adelante para publicitar y dar acceso a sus datos en formatos diversos a los otros actores de la sociedad. Primero, esto es importante porque la investigación, para ser replicable y verificada, tiene que ser transparente y abierta. Los beneficios son múltiples, no solamente en la transparencia de las múltiples investigaciones. Los estudiantes, como usuarios fundamentales de los datos, pueden ser beneficiarios directos de una posible apertura de datos de las universidades. Por otro lado, esta apertura puede ser un detonador y descentralizador de innovación, al dar acceso a datos de proyectos de investigaciones finalizadas, se facilita que otros investigadores puedan dar seguimiento a las preguntas subsecuentes de las investigaciones.

1.3.3 IMPULSAR EL “EMPREDIMIENTO DE DATOS”, A TRAVÉS DEL APOYO A LA INNOVACIÓN Y LA INCUBACIÓN DE NUEVAS EMPRESAS

El reciente caso de éxito de ODINE (Open Data Incubator Europe)³², así como el conocido surgimiento de emprendimientos de datos en Reino Unido y Estados Unidos han demostrado la oportunidad de generar valor, escalabilidad y ganancias con los emprendimientos de datos abiertos. En Latinoamérica también existen casos de emprendimientos de datos exitosos que han recibido inversión extranjera y han crecido en los últimos 5 años, tal es el caso de México con compañías como OPI, Data4, Atlantia Search, entre otras. Para promover un mejor entendimiento de esta dinámica, el GovLab lanzó el proyecto Open Data 100 México, que se puede consultar en línea. El objetivo de este estudio es triple: ”(1) Proveer una base para evaluar el valor económico y social de los datos gubernamentales abiertos; (2) Estimular el desarrollo de nuevas compañías que trabajan con Datos Abiertos; y (3) Facilitar diálogo entre el gobierno y el sector privado sobre cómo mejorar la utilidad de los datos gubernamentales abiertos”.

Definimos el “emprendimiento de datos” como emprendimientos de pequeño y mediano tamaño que hacen parte del ecosistema de Datos Abiertos y usan, generan o facilitan la generación de datos abiertos para construir negocios sostenibles que generen impacto económico, social y/o ambiental. De acuerdo con las últimas tendencias e investigaciones acerca de datos abiertos, se identifican dos buenas prácticas vinculadas a este tipo de datos que atienden, por un lado, la necesidad de crear o fortalecer empresas tecnológicas que desarrollen herramientas, plataformas, o servicios relacionados a Big Data (empresas que tradicionalmente entran en la categoría de TIC); y por el otro, fomentar a determinados sectores, industrias o empresas existentes para que usen los beneficios de los datos, en particular de los datos abiertos para la optimización de sus ventas, decisiones o procesos.

En Reino Unido, el gobierno se ha encargado de ofrecer datos abiertos gubernamentales de la más alta calidad a través de data.gov.uk. El Open Data Institute (ODI) en su investigación Open Data Means Business, ha analizado 270 empresas en el Reino Unido que usan, producen o invierten en datos abiertos como parte de su estrategia de negocio. Estas empresas (llamadas también “compañías de datos abiertos”) facturan más de 110 millones de dólares al año y emplean a más de 500 mil personas. La mayoría (70%) de las empresas analizadas usan datos abiertos proporcionados por el gobierno, aunque también usan datos abiertos de otras fuentes como Open Corporates. El sector privado está conformado por un gran número de pequeñas, medianas y grandes empresas, incluyendo grandes consorcios o nuevas empresas de tecnología y telecomunicaciones. Un análisis realizado por la firma de auditoría RSM encontró que para finales de febrero de 2017, el número de negocios de tecnología operando en el país había alcanzado los 5,995, con las concentraciones más altas de crecimiento de compañías tecnológicas en Londres.

Existen muchos tipos de compañías: regionalmente la mitad se encuentra localizada en la zona conurbada de Londres, una tercera parte tienen más de 10 años de antigüedad (lo cual no las califica como startups), el 70 por ciento tiene menos de 10 empleados, y el 54 por ciento se identifica como una empresa del sector TIC, aunque se registraron 12 sectores dentro de este tipo de compañías. Son empresas muy variadas que usan datos abiertos para encontrar un nicho de mercado, para desarrollar nuevos modelos de negocios o tomar decisiones sobre dónde poner una nueva tienda.

Algunos ejemplos de estas empresas son: Mastodon C, especialistas en Big Data que desarrollaron una plataforma open source para toma de decisiones en ciudades; Carbón Culture, una plataforma comunitaria diseñada para ayudar a las personas a utilizar los recursos de manera más eficiente, realizaron un proyecto de datos abiertos para ahorrar carbono, energía, agua y dinero; y Geolytix que ofrece servicios de geolocalización a retail, bancos y desarrolladores inmobiliarios utilizando datos abiertos.

Otro de los ejes relevantes sobre la relación del gobierno con las empresas TIC que usan datos abiertos es en la contratación de éstas últimas, es decir, el gobierno como cliente de los emprendedores de datos. Los gobiernos son usuarios de datos abiertos, por ejemplo para mejorar los servicios públicos y la eficacia gubernamental. Es posible, que las capacidades tecnológicas y de recursos humanos del gobierno no sean suficientes para generar, por ejemplo, un análisis con datos públicos. Por ello, el gobierno debe facilitar la contratación de empresas de datos, de forma competitiva y transparente. Los tipos de servicios de contratación son muy variados, puede ser desde una etapa inicial para la implementación de una estrategia nacional de datos abiertos, o en una etapa más consolidada donde los mismos datos que genera el gobierno se quieran usar para desarrollar modelos estadísticos avanzados.

Financiamiento de proyectos tecnológicos e incubación de nuevas empresas

Una de las medidas que puede tomar el estado para impulsar el “emprendimiento de datos”, es apoyar y financiar la incubación de nuevas empresas. Por ejemplo, en Reino Unido, se ha creado una agencia de innovación (Innovate UK), un organismo público ejecutivo no departamental, patrocinado por el Departamento de Negocios, Energía y Estrategia Industrial. Innovate UK se ha encargado de financiar proyectos tecnológicos que usan datos abiertos. Entre sus iniciativas se encuentra el [Programa Catapult](#), centros de innovación donde las empresas de todos los tamaños son apoyadas para transformar sus negocios a través la innovación digital. Para las Pymes, el enfoque está en “acelerar el crecimiento y el desarrollo sostenibles en las innovaciones digitales para fortalecer la economía del Reino Unido”. Para las grandes empresas, Digital Catapult busca “reforzar la cultura de la innovación e impulsa la colaboración con los innovadores digitales de Reino Unido, expertos digitales e investigación aplicada”.

Del mismo modo, pero desde el Open Data Institute (ODI), se desarrolló un programa en el cual se promueven, capacitan y conectan startups que utilizan datos abiertos llamado ODI Startup Accelerator. Su objetivo es “ayudar a las startups a innovar con datos, creando impacto económico, social y ambiental”. A lo largo de sus tres años de vida, este programa “ha ayudado a crear cientos de empleos” y se está desarrollando en dos ciudades del Reino Unido: Leeds y Londres.

En esta parte “Datos abiertos para impulsar la economía”, exploramos el rol del Estado, como poseedor y publicador de datos abiertos, y en la promoción de los mismos para impulsar la industria y el crecimiento económico. Como vimos anteriormente, poner a disposición datos abiertos de calidad y en grande cantidad es crear oportunidades de innovación, en particular para el sector privado. La promoción de una mayor apertura de los datos públicos, a diferentes niveles del gobierno, y de los datos en otros sectores (en particular empresas y academia), son medidas que pueden impulsar la economía. Asimismo, la multiplicación de “emprendimientos de datos” es fuente de crecimiento económico y puede ser promovida (por el gobierno) facilitando la innovación y la incubación de nuevas empresas.

2. CONSTRUIR Y ASEGURAR CAPACIDADES HUMANAS Y DE CAPITAL

2.1. MEJORAR LA EDUCACIÓN Y LAS COMPETENCIAS EN ANÁLISIS DE DATOS

Las “brechas de conocimiento pueden retardar el crecimiento de las empresas y pueden poner en peligro la empleabilidad de los trabajadores”³³ con implicaciones directas para el desarrollo económico y social de los países. Estas brechas se pueden profundizar aún más si, tanto la oferta como la demanda de educación y capacitación, no logran adaptarse a las transformaciones globales, como los cambios demográficos, los mercados globalizados, la innovación tecnológica y el cambio climático. Otro factor importante, en especial en países emergentes, es la carencia de educación de calidad resultando en una carencia en conocimientos y habilidades, concentrada en particular en la economía informal, obstaculizando el desarrollo de un ecosistema saludable para la industria.

En esta sección se analizan dos periodos de la formación de los individuos de forma separada: por una parte la estrategia de gobierno en educación primaria y secundaria, y por otra, la estrategia de gobierno a nivel de educación superior, laboral y ciudadano.

El enfoque “mente y mano” del MIT

En muchos países el aprendizaje de las habilidades básicas de datos (por ejemplo, Costa Rica, en toda la UE, India, China) incluye aprender inglés como segundo idioma. El enfoque práctico “constructivista” de usar datos para construir y controlar constructos físicos debe continuar a través de la universidad y el posgrado. Normalmente, esto implica construir sensores y sistemas informáticos (por ejemplo, controladores Arduino) para construir dispositivos físicos simples. Es ilustrativo que el lema del MIT es “mente y mano”, reflejando la necesidad de aprender a aplicar el conocimiento en situaciones reales para dominar el conocimiento.

Fuente: Alex Pentland

2.1.1 UNA FORMACIÓN EN DATOS Y TIC EN EDUCACIÓN PRIMARIA Y SECUNDARIA

“La lección de muchos países es que las habilidades básicas de datos deben volverse universales, comenzando en la escuela primaria. Las habilidades básicas incluyen el uso de datos y herramientas simples como hojas de cálculo para realizar proyectos escolares o como la construcción de máquinas simples”³⁴. Dos países que destacan por de sus esfuerzos en esta área son India y Singapur.

El primer esfuerzo organizado de India hacia la alfabetización en computación en los colegios fue el Proyecto CLASS (por sus siglas en inglés, Computer Literacy and Studies in Schools Project) iniciado en 1984. El objetivo del proyecto incluyó “la desmitificación de los computadores y su experiencia práctica. En 1993 el proyecto alcanzó a cubrir 2598 colegios operados por el Gobierno Nacional.”³⁵ Una vez el proyecto fue evaluado a través del Primer Estudio de la Computación en la Educación de India en 1993 (CompEd), se evidenció que, a pesar de que CLASS había alcanzado algunos avances en la alfabetización informática en el país (incremento en los programas de informática en los colegios), existía un número limitado de computadores para enseñar, una estructura organizacional ineficiente y no había un mecanismo de seguimiento y evaluación apropiado. Todo esto, sumado a cambios políticos, hizo que con el tiempo, el proyecto dejara de funcionar.

En 1998, el gobierno de India adoptó el Plan de Acción de Tecnologías de la Información (IT Action Plan) para la promoción de las TIC en los colegios. Este plan incluyó la implementación de una serie de acciones en algunos colegios pilotos denominados Smart Schools. El objetivo era mejorar la calidad de la educación a través de herramientas tecnológicas. Después de la publicación del IT Plan de Acción, en 2001, “el Gobierno lanzó un nuevo programa informático para los colegios, New CLASS, que tenía un espectro de acción más amplio que el primer programa CLASS, ya que no estaba restringido a los colegios operados por el Gobierno Nacional. El Gobierno se puso la meta de introducir alfabetización informática en 10.000 colegios, enseñanza asistida por computador en 1000 colegios y aprendizaje basado en la computadora en 100 colegios Smart. La lógica del Programa era que cada colegio Smart actuaba como mentor de otros colegios y adoptaba por lo menos 10 colegios de la provincia (estado), poniendo a disposición su infraestructura informática y capacitando a otros profesores en actividades pedagógicas de informática. El Gobierno central financiaba 1000 colegios mentores para necesidades de hardware y, para software y material didáctico, los colegios buscan fondos a través de costos adicionales para los estudiantes, donaciones individuales y contribuciones de la comunidad. Los gobiernos locales estaban obligados a asumir 25 por ciento del gasto total en hardware y software en los colegios bajo su jurisdicción.”³⁶ Esta serie de políticas e intervenciones del Gobierno en la educación primaria y secundaria hicieron posible que un número elevado de la población adquiriera habilidades informáticas, haciendo que las oportunidades de formación y aprendizaje (formal o no) pudieran tener lugar.

En Singapur uno de los elementos claves del gobierno actual para estimular el sector de Big Data es la construcción de capacidades y competencias informáticas a partir de una serie de programas y proyectos, que forman parte de Smart Nation Initiative (SIN), que incluye: IMDA Playmaker Program que expone a niños de manera interactiva a la tecnología, y el movimiento Code@SG que introduce pensamiento informático y de programación en estudiantes.

Este enfoque de promover el desarrollo de capacidades desde la educación primaria y secundaria tanto en India como en Singapur permite despertar talentos desde una edad temprana. Si bien la educación asistida por computador no implica, de por sí, un mayor éxito en el proceso aprendizaje³⁷ (por lo que su implementación debe ser cuidadosamente diseñada), familiarizar a las personas con “lo digital” facilita su inserción e inmersión en la economía digital. Esto podría explicar el por qué hay un alto porcentaje de estudiantes de secundaria que están interesados en estudiar ingeniería en India (“en 2015 cerca del 80% de los estudiantes indios están interesados en ingeniería ya que creen que la profesión ofrece un alto potencial de ingresos y la oportunidad de contribuir a las innovaciones”³⁸).

2.1.2 UNA FORMACIÓN EN DATOS Y TIC EN EDUCACIÓN SUPERIOR, EN EL MUNDO PROFESIONAL Y CIUDADANO

Es indudable que la formación no sólo debe tener lugar durante el periodo escolar, sino también llevarse a cabo durante la vida universitaria y laboral de los ciudadanos. El sector privado puede jugar un rol fundamental para que la formación sea adaptada a las necesidades de la industria y esto, a su vez, impacta positivamente en el empleo.

Regulation”, International Data Privacy Law, 2017. University of Oxford, Oxford Internet Institute. (2016)

²⁹ O’Neil, C. (2016). Weapons of math destruction: how big data increases inequality and threatens democracy. New York: Crown.

³⁰ Future of Privacy Forum, Understanding corporate data sharing decisions: practices, challenges, and opportunities for sharing corporate data with researchers, Noviembre 2017

³¹ www.opalproject.org

³² De acuerdo con la evaluación de impacto del programa publicada en 2017, “ODINE logró inspirar y promover una gama de nuevas ideas de negocios que aprovecharon el valor agregado de Open Data. Cada una de las 57 startups y Pymes financiadas utiliza dos o más tipos de datos abiertos, en particular datos geoespaciales/de mapeo y datos ambientales”. Como resultado, se estima que ODINE tuvo un impacto relevante en las perspectivas de crecimiento de la compañía, estimado a “110 € M de ingresos acumulados en el período 2016-2020, más 784 puestos de trabajo creados”. Para más información: Impact Assessment of ODINE Programme by IDC, July 2017. Disponible: https://opendataincubator.eu/files/2017/07/ODINE_Final-report_2.0.pdf

³³ International Labour Organisation (ILO). A skilled workforce for strong, sustainable and balanced growth: a G20 training strategy. 2010.

Para caracterizar el mercado laboral de la industria de Big Data se presenta en la tabla los resultados de la encuesta de Stack Overflow de 2017³⁹ (comunidad de más de 7 millones de programadores a nivel global) y líneas de acción que los gobiernos pueden ejercer para potenciar este mercado.

Tabla 1: Resultados de la encuesta de Stack Overflow y líneas de acción.

Resultado	Línea de acción
Predominio masculino - sólo el 7,6% de los encuestados son mujeres	Promover el desarrollo prioritario de capacidades de mujeres, así como el entendimiento y la eliminación de barreras de entrada de las mujeres a este mercado.
Preferencia por el trabajo remoto - 53,3% de los encuestados valoran el hecho de poder trabajar remotamente	Permitir formas flexibles y creativas de trabajo, que permitan la confianza, la motivación y la obtención de resultados
Mayoría autodidacta - el 90% de los desarrolladores se consideran a sí mismos autodidáctas	Ofrecer alternativas de aprendizaje a través de por ejemplo cursos en línea o campamentos de programación
Oferta desequilibrada - los perfiles de los encuestados son: científico de datos (8,4%), especialista en aprendizaje de máquinas (3,8%), desarrolladores Web (72,6%), de aplicaciones para ordenadores (28,9%) y para teléfonos móviles (23%)	Asegurar que para todas las especialidades existan opciones de formación

Un marco para desarrollar habilidades que correspondan a las necesidades de las empresas que demandan talento en Big Data, debe considerar la necesidad de contar con una fuerza laboral con conocimiento amplio en múltiples temas, y no sólo Big Data. Todo esto, porque “el problema de lejos, es encontrar gente con la combinación correcta de habilidades: científicos de datos que combinan sus habilidades técnicas, conocimientos analíticos y sobre la industria, y el sentido de negocios y habilidades blandas para hacer que los datos generen valor”⁴⁰. Modelos innovadores de enseñanza, por ejemplo vinculados al desarrollo empresarial, como prácticas de estudiantes con las empresas durante sus estudios o el establecimiento de grupos de trabajo entre estudiantes para resolver desafíos en Big Data de las empresas son algunas de las iniciativas que apuntan a resolver este problema.

En India, se destaca el desarrollo de un número elevado de programas de educación superior en análisis de datos y en negocios. “Hoy, India ha emergido como un jugador clave en el sector global de análisis de datos, particularmente en la provisión de servicios Business to Consumer (B2C). La rápida transición de TI a las tecnologías de última generación como Big Data y Data Analytics pueden atribuirse en parte a los institutos públicos de educación superior como el Instituto Indio de Tecnología (IITs) y el Instituto Indio de Negocios (IIMs)”⁴¹ y también a institutos privados como el NIIT. El aumento de la oferta de programas de Data Analytics en estos institutos, entre otros, contribuyó a asegurar que exista mano de obra calificada disponible en el mercado. Al mismo tiempo, India, es “el mayor proveedor de programas de educación en negocios en el mundo buena parte de ellos, con pensum estandarizados. Sin embargo, existe una gran brecha entre las mejores escuelas de negocios, como el Indian Institute of Management (IIM-A) calificado como MBA número uno en India por The Economist ranking of business schools y, otros programas menos conocidos en India.”⁴² No obstante, de acuerdo con diversos estudios⁴³, estos esfuerzos en educación terciaria en conocimientos de TIC (incluyendo datos) y negocios, hacen que en India la mano de obra india sea más competitiva.

Otro aspecto a resaltar de India en formación de capacidades digitales y análisis de datos, es el programa del gobierno actual de India Digital (Digital India)⁴⁴ cuyo impacto podría llegar a un 25 a 30 % del PIB en 2025 según Deloitte⁴⁵. Se puso en marcha en 2015⁴⁶ y el programa busca “transformar el país en una sociedad digital empoderada y una economía del conocimiento”. Dos de sus tres áreas-visión tienen estrecha relación con las capacidades en análisis de datos: gobierno y servicios a medida

³⁴ Fuente: Alex Pentland

³⁵ Aggarwal, Devi Dayal. History and development of elementary education in India, vol. 1. Sarup & Sons, 2002.

³⁶ T.J. Plomp, R.Anderson, N. Law and A. Quale. International Association for the Evaluation of Educational Achievement (IEA) Cross national information and communication technologies policies and Practices in Education. Chapter 19.P.3. Formation Age Publishing- IAP.

³⁷ Leigh Linden, Complement or Substitute? The Effect of Technology on Student Achievement in India, Columbia University, 2008.

³⁸ The Times of India, Most Indians prefer engineering as a career option: Report, The Times of India, 2015.

³⁹ Muestra de 64.000 personas (de las cuales, cerca de 200 eran de Colombia), de diferentes profesiones como científicos de datos, ingenieros informáticos, analistas e ingenieros de datos, especialistas en aprendizaje de máquinas, etc.. La encuesta pregunta por sus tecnologías favoritas, sus costumbres para codificar, sus preferencias de trabajo, las maneras cómo aprenden, entre otros.

⁴⁰ Obijiofor, Levi. New Technologies in Developing Societies. From Theory to Practice. Palgrave Studies in Communication for Social Change. 2015. Averou, Chrisanthi. The link between ICT and economic growth in the discourse of development. London School of Economics. Gollakota Kamala, Pick, James B and Sathyapriya. Using technology to alleviate poverty: use and acceptance of telecenters in rural India. Volume 18. Issue 3. 2012.

de la demanda, y empoderamiento digital de los ciudadanos. Estas dos áreas se dividen a su vez en seis pilares/programas: gobierno electrónico (e-government), información para todos, e-kranti (provisión electrónica de servicios), manufactura de artículos electrónicos, TIC para empleos y programas de semilla. Cada una de estas áreas se desarrolla de manera independiente y coordinada por diferentes departamentos y ministerios. En el marco de cada pilar se desarrollan diferentes proyectos e iniciativas. Una de ellas es el programa Skill India que busca formar 400 millones de personas en diferentes temáticas (incluyendo alfabetización digital y proyectos especiales), en el período de 2015 a 2022. Todavía es muy prematuro para decir que Digital India está teniendo éxito. Lo que es interesante resaltar, es la promoción transversal de TIC en distintos espectros de la sociedad y con el propósito de generar alternativas económicas para las personas y las empresas.

Por otra parte, el gobierno y el sector privado han iniciado una serie de iniciativas para crear conciencia y aumentar el conocimiento digital en India, “la Misión de Alfabetización Digital, (National Digital Literacy Mission- NDLM por sus siglas en inglés) creada en 2014 es una de ellas y ha beneficiado 8.5 millones de personas (se ha desarrollado a través de Alianzas Público-Privadas). El Gobierno busca ampliarla para cubrir un porcentaje mayor de la población en zonas rurales. Otra iniciativa de formación se dio en diciembre de 2014 en el marco de una alianza entre el gobierno e Intel, donde a través de una aplicación en 5 diferentes dialectos, se capacita a personas en temas digitales, de inclusión financiera y salud⁴⁷.” El departamento de Ciencia y Tecnología de India, a su vez, a través de su iniciativa Big Data ha previsto invertir en el desarrollo de capacidades y habilidades a largo plazo y crear alianzas entre la industria y la academia para fortalecer las universidades y desarrollar currículos de formación en análisis de datos.

En cuanto a Singapur, otros dos programas que también forman parte de la Smart Nation Initiative (ver sección 2.1.1) son Skills Future, que desarrolla capacidades en TIC para el trabajo y el Smart Nation Fellowship Program que invita a científicos de datos e ingenieros a participar de la solución de desafíos nacionales o de otros países en diferentes áreas.

Así también, la orientación del Gobierno de Singapur en cuanto al desarrollo de acciones que permitan contar con ciudadanos y trabajadores con conocimientos y habilidades en tecnologías de la información y las comunicaciones, es relevante, en la medida en que establece programas diferenciados para todas las personas de la sociedad. La Agencia del Gobierno IDA, primero identificó seis áreas de necesidades de formación en donde toda la población del país requería capacitarse y formuló los siguientes programas⁴⁸:

1. Programa Nacional de Alfabetización en Tecnologías de la Información (NITLP): enfocado en no especialistas en informática (trabajadores, amas de casa y adultos mayores) para dotarlos de habilidades básicas de computación e internet, que les permita mejorar su calidad de vida y su empleabilidad.
2. Programa de Competencias de Información y Comunicación (ICP): dirigido a entrenar trabajadores en competencias esenciales de TIC para ser más eficientes en sus lugares de trabajo.
3. Programa E-Business Savviness (EBSP): busca desarrollar un grupo de mano de obra calificada con habilidades y conocimientos en negocios electrónicos.
4. Programa Estratégico de Conversión de la Mano de Obra (SMCP): para convertir a los profesionales no TIC en profesionales TIC.
5. Programa de Recursos de TIC (CITREP): apunta a mejorar las habilidades de profesionales en áreas críticas y/o emergentes.
6. Programa de Entrenamiento y Logros para la Capacitación en TIC: busca fortalecer la industria a través de capacitaciones mucho más especializadas.

Posterior a la identificación de estos seis niveles de capacitación, el gobierno trabaja con el sector privado, la comunidad y otras instituciones públicas para implementar cada uno de los programas.

Si consideramos el caso del Gobierno Australiano, la alfabetización en datos (“data literacy”) ha sido promovida al interior de la administración pública. Así, después de encontrar que el país tenía un escaso conocimiento en materia de datos, y que esto limitaba su capacidad de sacar el máximo provecho a los datos disponibles (Informe de Gestión de Datos del Sector Público en 2015), el Gobierno australiano estableció un Marco de Capacidades de Datos del Servicio Público Australiano, para definir una hoja de ruta para crear capacidades de análisis y uso de datos en la administración pública. Cuatro componentes comprenden ese Marco de desarrollo de habilidades en Datos: (1) el programa de becas en datos; (2) cursos universitarios; (3) el programa alfabetización de datos en la administración pública y (4) las asociaciones/alianzas para la formación/entrenamiento en datos. La idea es que los altos ejecutivos de la administración animan a los empleados públicos a aprovechar estas oportunidades de aprendizaje y desarrollo. Esta política además de contribuir al mejoramiento de las tareas del gobierno en el sector público, tiene un doble enfoque positivo en el sector privado: ayuda a sentar las bases para que los procesos entre lo público y privado puedan ser más eficientes y más orientados a los datos, y transmite coherencia, compromiso y buen ejemplo ante otros actores.

⁴¹ Más información sobre Digital India puede ser consultada en: <http://www.digitalindia.gov.in/content/introduction>

⁴² Deloitte y ASSOCHAM India, “Digital India: Unlocking the trillion dollar opportunity,” 2016

En síntesis, desde las experiencias de India, Singapur y Australia la estrategia política debe apuntar a una formación en datos y TIC presente a nivel superior, profesional y ciudadano que permita la generación de mano de obra calificada y que a su vez se adapte a dinámicas propias del mercado laboral Big Data e interfiera en ellas. Por otro lado, se necesita que la formación empodere a los ciudadanos mediante la alfabetización digital para no especialistas y así facilite la transformación hacia un gobierno electrónico. Además, la formación en datos y TIC debe incluir habilidades blandas y de negocios que permitan potenciar al sector privado e integrarlo tanto en la articulación y el diseño del currículo universitario, como en la formación de estudiantes y trabajadores.

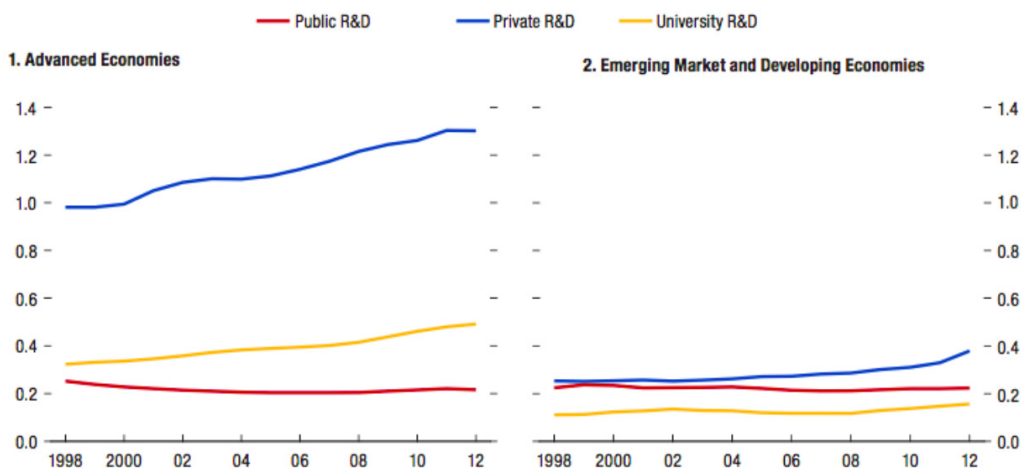
2.2. INVESTIGACIÓN Y DESARROLLO PARA BIG DATA

En esta sección se presenta los principales elementos a tener en cuenta en el diseño de una estrategia que fomente la investigación y desarrollo de empresas entorno a los temas de Big Data. Para ello, se comienza detallando la diferencia en el nivel de gasto general en investigación y desarrollo que realiza el sector público y privado. Luego se listan algunas experiencias de otros países para la promoción de la investigación y desarrollo en torno al Big Data. Por último, se presentan las conclusiones de esta sección.

La implementación de políticas gubernamentales para apoyar la investigación y el desarrollo del sector de Big Data, sostiene un artículo publicado por el FMI en 2016, es fundamental para el crecimiento de la economía global en general, en particular considerando el lento crecimiento en otros sectores desde la recesión económica de 2008. El apoyo de la investigación y desarrollo de Big Data también tiene un potencial significativo. Como el informe del FMI describe, comparado con las economías avanzadas, las emergentes tienen proporcionalmente menos investigadores, patentes y publicaciones, dejando un espacio significativo para el crecimiento.⁴⁹ Además, como enseña la figura a continuación, hay una brecha aún mayor entre la I+D financiada con fondos privados y la I+D pública comparativamente más pequeña. Por otra parte, la I+D pública por lo general produce una menor tasa de rendimiento y, por lo tanto, la promoción de la I+D de las empresas privadas es más beneficiosa.⁵⁰

Figura 3: Gastos de investigación y desarrollo, 1998-2012

Private R&D and university R&D expenditures are significantly higher in advanced economies than in emerging market and developing economies. R&D expenditure is also increasing more rapidly in advanced economies. Public R&D expenditure is similar across groups of economies.



Source: United Nations.

⁴³ Un esfuerzo anterior realizado por el Gobierno de India fue el Programa Nacional de Desarrollo de Habilidades puesto en funcionamiento en 2009. El objetivo era empoderar a todos los ciudadanos a través de nuevas habilidades, conocimientos y calificaciones con reconocimiento internacional para darles acceso a empleo decente y para promover un crecimiento nacional inclusivo. Así también, aumentar la capacidad de entrenamiento vocacional.

⁴⁴ Obijiofor, Levi. *New Technologies in Developing Societies. From Theory to Practice*. Palgrave Studies in Communication for Social Change. 2015. Averou, Chrisanthi. *The link between ICT and economic growth in the discourse of development*. London School of Economics. Gollakota Kamala, Pick, James B and Sathyapriya. *Using technology to alleviate poverty: use and acceptance of telecenters in rural India*. Volume 18. Issue 3. 2012.

⁴⁵ Nagy K Hanna and Peter T Knight, *National strategies to harness information technology: Seeking transformation in Singapore, Finland, the Philippines, and South Africa*. Springer, 2011.

⁴⁶ World Bank, *World Development Indicators*. Disponible: data.worldbank.org/data-catalog/world-development-indicators

⁴⁷ Como señala el documento del FMI, existen dificultades metodológicas para medir el rendimiento de la investigación científica básica a causa de los largos retrasos y las limitaciones de datos, especialmente a nivel macro.

El progreso tecnológico de una economía depende de las instituciones y de las políticas públicas. De acuerdo con lo mencionado en la sección 1, la inversión en educación y otros aspectos del capital humano tienen un impacto significativo en el desarrollo de una economía de Big Data, pero son apenas una parte de una recomendación más amplia de política pública.

Para garantizar que los fondos públicos se asignen de la manera más eficiente posible para generar crecimiento e innovación en el ecosistema de Big Data, la focalización del gasto público en el ecosistema de Big Data tiene que ir acompañado de un análisis que estime la competitividad del país en cada una de las áreas de la “cadena de valor” del Big Data. Esta cadena, tal como la define el Prof. Pentland, indica la prioridad o la importancia de varios enlaces dentro del ecosistema. La cadena consiste de: acceso y recopilación, integración, limpieza, control y consolidación, protección, análisis, y aprovechamiento operativo de la información. Al identificar las etapas en las cuales un país se encuentra más avanzado, se puede identificar las etapas menos desarrolladas y dirigir más recursos a su crecimiento. Algunas preguntas relevantes que una administración puede considerar pueden incluir: ¿Qué porcentaje de los emprendimientos de datos, de las oficinas de estadística o del mismo gobierno, se enfocan en suministrar este tipo de servicios de datos? ¿Qué tan eficientes y manejables son los datos y cómo las instituciones que los administran pueden garantizar que dichos datos puedan completar su proceso?

La política fiscal, a través de los incentivos tributarios puede ser un canal para promover la investigación en Big Data. “La política fiscal puede jugar un papel importante en estimular la innovación a través de sus efectos en la investigación y desarrollo, emprendimiento, y transferencia de tecnología.”⁵¹ Por ejemplo, darle beneficios tributarios a las empresas que logran generar patentes de productos relacionados a Big Data. Sin embargo, lo anterior puede llevar a que las empresas más grandes sean las beneficiadas, dado que por lo general las empresas más grandes son las que concentran el mayor número de patentes. Esto podría llevar a pensar no solo en el número de patentes sino en el tamaño de la empresa y año de constitución. La promoción de adopción de nueva tecnología también es un elemento importante. Los altos costos para adoptar nueva tecnología es una barrera que puede ser atendida por la política pública. Los aranceles asociados al hardware o infraestructura necesaria para crear ecosistemas de Big Data deberían ser bajos o nulos.

Con respecto a Latinoamérica, la CEPAL ha dicho que existen tres niveles de intervención gubernamental cuando se habla de innovación en las pequeñas y medianas empresas. En el primer nivel se pueden encontrar los instrumentos del sistema formal de ciencia y tecnología y que están orientados a impulsar los procesos de innovación asociados a los gastos de Investigación y Desarrollo; normalmente son las empresas tecnológicas medianas las que responde a este tipo de estímulo. En segundo lugar, están los instrumentos indirectos o subsidios a la demanda. Por último, en el tercer nivel están las acciones donde el gobierno tiene un rol más activo: desde la creación de centros tecnológicos, los centros de extensionismo tecnológico relacionados con las universidades, los centros públicos, hasta la creación de incubadoras de emprendimientos.⁵²

Sin embargo, la mayor parte de recursos en los países de la región son bolsas de fondos de servicio universal o de acceso universal, que buscan en primera instancia en cerrar la brecha. El objetivo de este tipo de recursos es extender los servicios TIC a aquellas regiones o aquellos segmentos de poblaciones económicamente no viables para la iniciativa privada. Estos han funcionado para distribuir redes de acceso de telefonía fija y móvil, banda ancha, entre otros. Es importante que en las zonas donde ya se cuenta con niveles de cobertura y tecnología deseable se pueda desarrollar programas en torno al uso del Big Data. Un ejemplo de ello, es promover la cultura de datos y la importancia de tener buena infraestructura tecnológica. Muchas veces los gobiernos invierten gran cantidad de recursos en servicios TIC y la población no entiende sus ventajas y potenciales usos.

Los fondos de servicio universal en Latinoamérica han logrado a través de impuestos específicos en el sector TIC o mediante presupuestos creados por decisión del Ejecutivo. Destaca la inversión pública en acceso universal de Brasil, con una media de alrededor de 25 dólares por persona al año.⁵³ Existen otros países, como Uruguay, donde no existe un fondo y este tipo de servicios los implementa directamente la oficina de telecomunicaciones del gobierno.

Para el caso de Estados Unidos, un país líder en la innovación y desarrollo de Big Data, ha expandido significativamente su programa de investigación y desarrollo en este campo en los últimos años. La Administración de Obama invirtió más de 200 millones de dólares en la Iniciativa de Investigación y Desarrollo de Big Data. A través de fondos nuevos y con la participación de recursos del presupuesto de seis agencias del gobierno se invirtió esta suma para mejorar la capacidad de extraer conocimiento e información del Big Data. A través de la Oficina de Política de Ciencia y Tecnología de la Casa Blanca (OSTP) y la Fundación Nacional de Ciencias (NSF), junto con los Institutos Nacionales de Salud (NIH), el Departamento de Defensa (DoD), la Agencia de Proyectos de Investigación Avanzada de Defensa (DARPA), la Oficina de Ciencia de la Energía (DoE) y el Servicio Geológico de los Estados Unidos (USGS, por sus siglas en inglés), se realizó este gasto histórico para promover el desarrollo de las tecnologías más avanzadas de Big Data. El objetivo era aprovechar estas tecnologías para acelerar el ritmo del descubrimiento en la ciencia y la ingeniería, fortalecer su seguridad nacional y transformar la enseñanza;

⁴⁸ Nagy K Hanna and Peter T Knight, *National strategies to harness information technology: Seeking transformation in Singapore, Finland, the Philippines, and South Africa*. Springer, 2011.

⁴⁹ World Bank, *World Development Indicators*. Disponible: data.worldbank.org/data-catalog/world-development-indicators

⁵⁰ Como señala el documento del FMI, existen dificultades metodológicas para medir el rendimiento de la investigación científica básica a causa de los largos retrasos y las limitaciones de datos, especialmente a nivel macro.

del mismo modo para ampliar la fuerza de trabajo necesaria para desarrollar y utilizar las tecnologías de Big Data. En la primera etapa se gastaron 25 millones por parte del NSF y NIH para investigación en bases de datos relacionadas con datos moleculares, celulares, electrofisiológicos, químicos, comportamientos, epidemiológicos, clínicos, y otros conjuntos de datos relacionados con la salud y las enfermedades. Posteriormente, el Departamento de Defensa lanzó “Data to Decisions” con una inversión de 250 millones de dólares anuales, acompañada de 60 millones de dólares exclusivamente para nuevos proyectos de investigación, en una serie de programas para apoyar la toma de decisiones en temas de seguridad.

Sin embargo, en Estados Unidos el gasto en este tipo de proyectos va más allá del presupuesto I+D. En 2015, la inversión en servicios de información y software fueron aproximadamente 1,500 millones de dólares. Por su parte, en cuanto almacenamiento de datos se utilizaron alrededor de 3,400 millones de dólares. Es importante destacar que desde la propuesta presupuestal de Obama en 2012 el gasto en investigación y desarrollo se enfocó en temas militares y de defensa. Ahora, en el plan presupuestal de la nueva administración, la investigación y desarrollo sólo aparece en temas biomédicos, y está calificado como de alta prioridad.

La perspectiva de la gran inversión en Big Data de Estados Unidos ha sido confrontada por el Instituto Brookings, que ha calificado sutilmente la inversión en datos y estadísticas del gobierno como pequeña.⁵⁴ De acuerdo con el artículo, los recortes discrecionales del presupuesto en este rubro le han quitado capacidad al gobierno para cumplir con sus deberes. Con una inversión poco representativa (tan sólo alrededor de un quinto del 1 por ciento del presupuesto federal) las oficinas de estadística (y otros organismos que ofrecen datos) proveen un servicio valioso para los negocios, las autoridades y las familias. Si bien no se ha medido el impacto de esta reducción de recursos para el desarrollo de Big Data, recortar el presupuesto para Big Data, resulta contradictorio para el desarrollo de un país, ya que no tener información para la toma de decisiones de política económica podría afectar el crecimiento de los distintos sectores de la economía.

Además de Estados Unidos, Reino Unido ha definido una inversión histórica de £4.7 billones en investigación y desarrollo (Declaración Otoño 2016) para promover la innovación. Otras organizaciones en ese país han empezado a desarrollar esfuerzos adicionales más focalizados en Big Data, como el Consejo de Investigación Social y Económico (ESRC por sus siglas en inglés), que siendo la organización más grande para financiamiento de investigación en iniciativas económicas y sociales, está financiando tres Centros de Investigación en Datos con un enfoque particular en negocios y datos de los gobiernos locales. Inversiones para crear las condiciones necesarias para desarrollar la industria de la inteligencia artificial están también dentro de los planes del gobierno inglés.

En conclusión, si se quiere desarrollar industria de alta tecnología, hay que promover la alta demanda de investigación científica y gasto en I +D, el nivel alto de innovación, fomentar la difusión rápida de innovación tecnológica, el alto nivel de empleo de personal científico y técnico, mitigar los altos riesgos de inversión y procesos rápidos de depreciación de la inversión, ayudar con la cooperación estratégica intensiva nacional o internacional con otras empresas de alta tecnología o centros de investigación, ayudar jurídica para la creación de patentes y licencias y finalmente, fomento de las competencia para la generación de nuevos negocios y comercialización

2.3. MODERNIZAR LA INFRAESTRUCTURA DE DATOS

Con el fin de entender las necesidades y tendencias del sector privado en temas de Big Data, Gartner ha venido realizando una encuesta a empresas líderes en temas de tecnología de la información. La última encuesta, desarrollada en 2016, tuvo un total de 199 participantes. Al hacer referencia a cuáles son los principales retos que una organización tiene frente al manejo de Big Data, la figura 4, muestra que determinar el valor de Big Data sigue siendo el principal reto en los últimos cuatro años. Sin embargo, este reto ha venido disminuyendo. Mientras que los retos asociados a habilidades y capacidades viene en aumento. En referencia al componente de infraestructura, 26% de los encuestados informó integración con infraestructura existente como un desafío clave para la adopción de Big Data. Por otro lado, el 19% de los encuestados respondió que infraestructura y arquitectura es un reto, sin embargo, este desafío cada vez es menor entre los encuestados.

Las entidades, tanto privadas como públicas, que trabajan con datos deben formalizar y consolidar distintas gamas de tecnologías y estructuras relacionadas con la información. Tanto la coordinación y conexión entre ellas es indispensable para un correcto funcionamiento de todo el sistema. A raíz de lo anterior, esta sección detalla y describe los distintos componentes que conforman la infraestructura física y tecnológica en torno al Big Data. Para ello, se comenzará explicando que se entiende por infraestructura de datos, para luego pasar a la descripción de los principales tipos de infraestructura y arquitectura de datos.

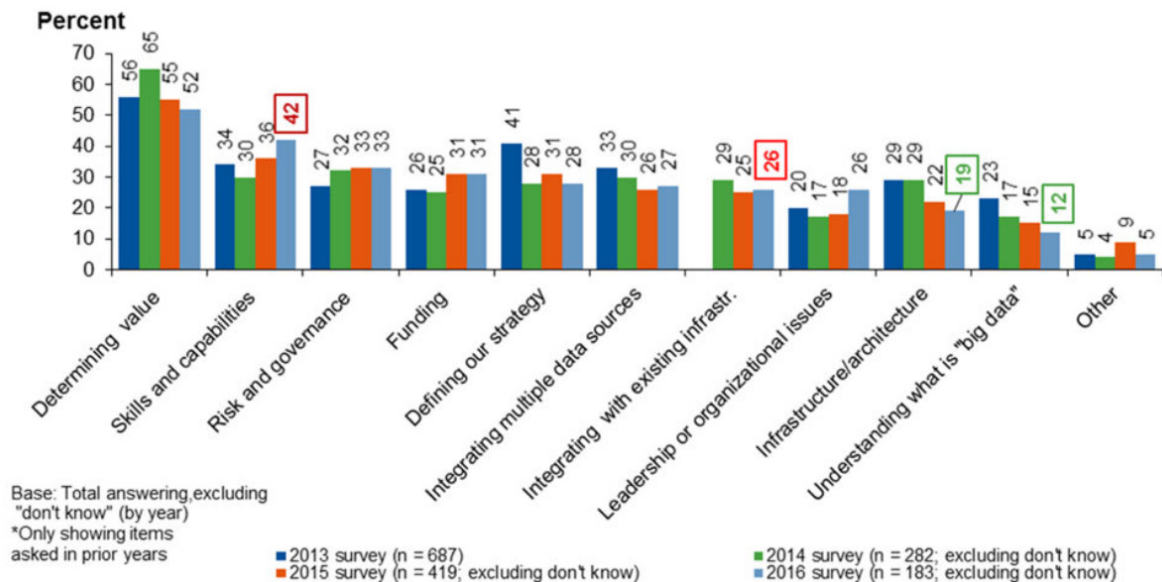
⁵¹ Ibid

⁵² Marco Dini, Giovanni Stumpo, and Italia Cooperazione Italiana. Políticas para la innovación en las pequeñas y medianas empresas en América Latina. 2011.

⁵³ Banco de Desarrollo de América Latina, Hacia la transformación digital de América Latina: las infraestructuras y los servicios TIC en la región. Banco de Desarrollo de América Latina, 2013.

⁵⁴ Diane Schanzenbach, America's small investment in government data has big payoffs, March 2017, <https://www.brookings.edu/opinions/americas-small-investment-in-government-data-has-big-payoffs/>. OECD Directorate for Science and Innovation. Data-driven innovation for growth and well-being. October, 2015.

Figura 4: ¿Cuáles son los principales retos de su organización con respecto a Big Data?



Fuente: Gartner, Septiembre 2016.

¿Qué se entiende por infraestructura de datos?

¿ Qué es la “infraestructura de datos”?

La infraestructura de datos incluye tecnología, procesos y organización. Particularmente, consiste en activos de datos, las organizaciones y sistemas que los operan y mantienen, y procedimientos y guías que describen cómo usar y administrar los datos. La infraestructura confiable de datos cuenta con financiamiento sostenible y supervisión que brinda orientación para maximizar el uso y el valor de los datos al satisfacer las necesidades de la sociedad.

Fuente: Alex Pentland

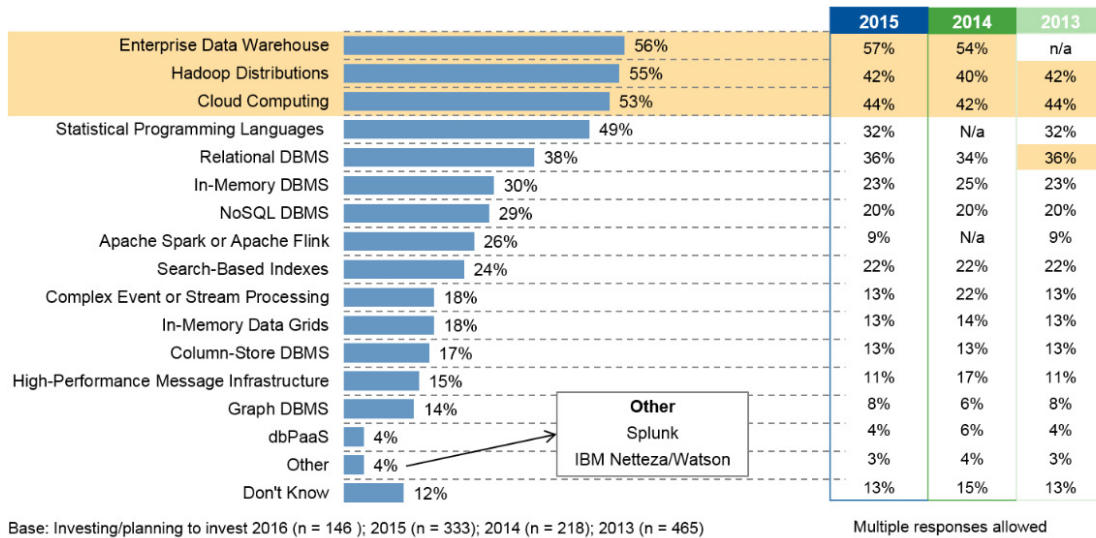
Lo importante sobre la infraestructura de datos es que los gobiernos puedan promover el desarrollo y adopción de tecnologías Big Data, computación en la nube y supercomputadoras, que les permita innovar y mejorar en el uso de la explotación de datos. Esto puede beneficiar a las organizaciones, en el sentido de permitirles anticiparse mejor, predecir y adaptar mejor sus productos y servicios a las necesidades de cada cliente o individuo. Estos beneficios se traducen, por una parte, en la reducción de costos operacionales para las empresas (mantenimiento predictivo de sus equipos, reducción del fraude, optimización de procesos clave) y por otra, en la creación de servicios innovadores (simplificación de las solicitudes para los ciudadanos a través de aplicaciones digitales, mejor comprensión de los problemas de los clientes, entre otros).

En general y como complemento a las tendencias presentadas por Gartner, según el reporte de la OCDE, “Data Driven Innovation for Growth and Well Being” de 2014⁵⁵, los desafíos más importantes a los que los países deben dar respuesta están del lado de la oferta, la demanda y la sociedad. En cuanto a la oferta de Big Data son los siguientes: (i) la inversión en banda ancha móvil (particularmente en regiones más remotas y menos desarrolladas) y en las barreras al libre flujo de datos, (ii) el acceso y la propiedad de los datos y los incentivos para compartirlos, esto para desarrollar y mantener los datos, y (iii) acceso a analítica y computación en la nube, incluyendo esquemas de licencias *open source*. Los desafíos del lado de la demanda están relacionados con la capacidad de sacar provecho de la innovación enfocada en datos (desarrollo de habilidades, cambios organizacionales y fortalecimiento al emprendimiento), mientras que del lado social hay aspectos que afectan tanto la oferta como la demanda y que incluyen la pérdida de autonomía y libertad, cambios estructurales en el mercado de trabajo, manejo de aspectos de seguridad y la concentración y la dominación del mercado.⁵⁶

De acuerdo con Gartner, en la figura 5 se puede apreciar que más del 50%, sombreado naranja, de las organizaciones usan o planean usar una bodega de datos, una distribución Hadoop y/o computación en la nube para obtener valor del Big Data. Lo interesante es ver cómo las bodegas de datos han sido la tecnología más utilizada para almacenar, analizar y administrar big data durante 2014, 2015 y 2016. Sin embargo, solo hasta el año 2016, Hadoop y la computación en la nube alcanzan el mismo uso de las bodegas de datos, como las tecnologías más utilizadas para Big Data. Lo anterior muestra que es importante

que las empresas que buscan generar valor a partir del Big Data deben contemplar el uso de distintas tecnologías para trabajar datos. Sin embargo, se evidenció que un reto que están viviendo las empresas es saber cómo integrar múltiples infraestructuras. Adicionalmente, las tecnologías de código abierto, como Hadoop, se han vuelto una opción atractiva para construir lagos de datos de cualquier tipo de información. Estas tecnologías de código abierto muchas veces ya vienen integradas en plataformas en la nube como Amazon Web Services, Google o Microsoft Azure, permitiendo el procesamiento de datos de alto rendimiento de forma rápida y de fácil uso.

Figura 5: ¿Qué tecnologías su organización usa o está planeando usar para obtener valor del Big Data?

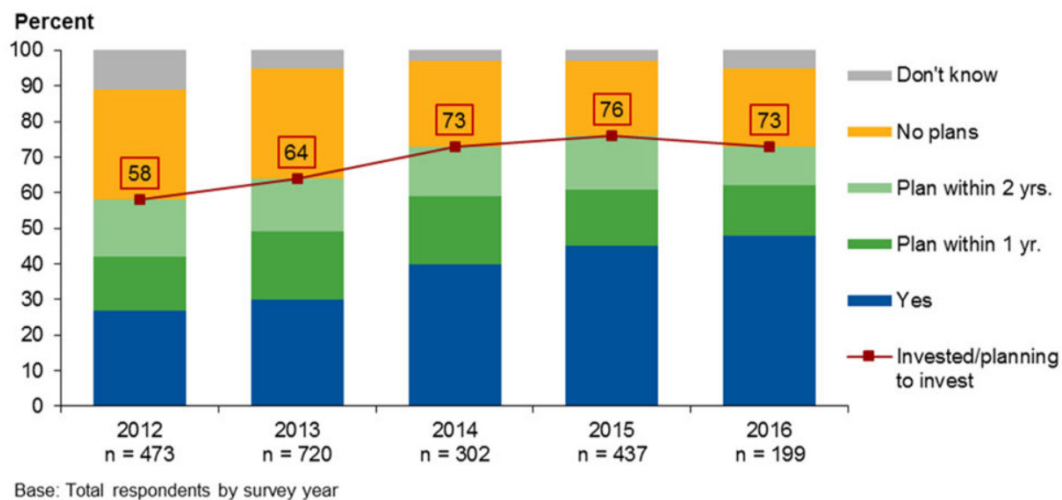


Fuente: Gartner, Septiembre 2016

La creación de la computación en la nube ha reducido parcialmente el problema de escala asociado con las infraestructuras TI, convirtiendo estos costos fijos en costos variables y permitiendo a las pequeñas empresas operar sin tener la propiedad de las infraestructuras físicas. Con compañías como Amazon, Google y Microsoft proveyendo algoritmos de machine learning como parte de sus servicios de computación en la nube, pequeñas compañías encuentran mucho más conveniente tener sus datos procesados y minados usando infraestructuras TI externas.⁵⁷ De hecho, Cisco proyectó que para 2020, 92 por ciento de todos los negocios procesarán su carga de trabajo a través de computación en la nube.⁵⁸

Lo anterior tiene efectos en el nivel de adopción de nuevas tecnologías o proyectos enfocados a Big Data, cada vez más empresas tienen presupuestado explorar y pilotear proyectos usando nuevas tecnologías de Big Data. Sin embargo, a pesar de que cada vez sea más baratas estas tecnologías, el desconocimiento de cómo usarlas o integrarlas a infraestructura vieja sigue siendo un reto y desafío para que este ritmo de adopción mantenga una tendencia creciente. Ver Figura 6.

Figura 6: Estado de adopción de Big Data en su organización



Fuente: Gartner, Septiembre 2016

Sin embargo, con el aumento del número de compañías que se vuelve dependiente de estas infraestructuras de pocos proveedores, estos últimos ganan mucho más poder en el ecosistema Big Data al acceder cada vez más, a volúmenes y variedades significativas de datos, que les permiten a ellos mejorar sus algoritmos y análisis de datos. Si las tendencias continúan, un problema de competitividad puede presentarse en el futuro, ya que muchos nuevos jugadores no estarán en capacidad de construir infraestructuras lo suficientemente fuertes, comparados con estos pocos proveedores. El verdadero desafío para los gobiernos será en este caso cómo regular y contener los efectos perversos de empresas proveedoras de infraestructura que como todos los oligopolios tienen un gran potencial de generar externalidades negativas para el buen funcionamiento del mercado, que pueden incluir guerras de precios (al favorecer “price stickiness”), retardar la innovación al no favorecer la entrada de nuevos competidores, complejizar la regulación, entre otros. Algunas de las herramientas con las que los gobiernos cuentan para enfrentar estos desafíos incluyen el diseño e implementación de políticas que permitan y garanticen el adecuado uso y tratamiento de la información de empresas en estas plataformas y que en caso de algún manejo inadecuado, las empresas dueñas de estas plataformas respondan adecuadamente a las demandas que hayan a lugar.

En resumen, los gobiernos deben considerar políticas orientadas a promover la demanda y adopción de tecnología e infraestructuras TI que favorezcan la innovación en Datos y al mismo tiempo establecer un marco regulatorio que asegure la protección de datos de los consumidores y que impida el desarrollo de prácticas anticompetitivas por parte de los proveedores. Los gobiernos pueden enfocar su acción en una serie de actividades que incentiven la competencia de mercado de productos y servicios para el aprovechamiento de datos, estas pueden incluir la sensibilización a las empresas sobre la importancia de contar con la tecnología adecuada para innovar con datos, un portal virtual con herramientas disponibles y el desarrollo de redes que permitan informarse bien sobre las infraestructuras tecnológicas disponibles.⁵⁹ Los gobiernos deben apoyar el desarrollo de la administración del riesgo digital en todos los ámbitos de la sociedad así como actuar si existen barreras regulatorias para adoptar nuevas tecnologías.

3. CONSTRUYENDO COMUNIDAD EN TORNO AL BIG DATA

En esta sección se presentan distintos mecanismos direccionados a la construcción y desarrollo de un ecosistema de Big Data. Esta comunidad incluye a los generadores, reguladores y usuarios de datos. En particular, se hace mayor énfasis en los productores y reguladores. Además, se presentan algunos conceptos y casos prácticos evidenciados en otros países.

3.1. FACILITAR ALIANZAS PÚBLICO-PRIVADAS

En un esfuerzo por crear una cultura que promueva y permita un entorno propicio para una economía sostenida de Big Data, un gobierno puede fomentar las Alianzas Público-Privadas (APP). Aunque no existe una definición globalmente aceptada de APP, éstas pueden ser entendidas como “un acuerdo entre dos o más organizaciones que les permite trabajar cooperativamente hacia objetivos compartidos o compatibles y, en la cual hay cierto grado de autoridad y responsabilidad compartida, inversiones conjuntas de recursos, toma de riesgo compartido, y beneficios mutuos”.⁶⁰ Para la OCDE, una APP es “un acuerdo entre el gobierno y uno o más socios privados (que pueden incluir a los operadores y los financiadores) según los cuales los socios privados entregan el servicio de tal manera que los objetivos de entrega del servicio del gobierno estén alineados con los objetivos de ganancias de los socios privados y donde la efectividad de la alineación depende de una transferencia de riesgo suficiente a los socios privados”⁶¹

Si bien, las APP están basadas en generar valor agregado mutuo, es el privado el que asume gran parte del riesgo y de la responsabilidad de la administración y remuneración que por lo general se encuentran vinculadas al rendimiento del proyecto mismo. Los beneficios de estas alianzas pueden tomar varias formas, como por ejemplo, de tipo financiero o material, o menos tangibles como know how o nuevo conocimiento. Del lado de los costos, hay de tipo unitario o recurrente.

⁵⁵ OECD Directorate for Science and Innovation. Data-driven innovation for growth and well-being. October, 2015.

⁵⁶ Según la OCDE una de las externalidades negativas que pueden presentarse es la creación o profundización de oligopolios. Los mercados basados en datos pueden conducir situaciones donde un ganador toma todo (a winner takes all). Si bien los datos pueden ser clave para el crecimiento y éxito de una empresa, a medida que exploramos, los datos pueden ser un factor crítico para el crecimiento y éxito de una empresa. En una de las fusiones que analizamos, la adquisición de Google de Waze, la incapacidad de Waze para lograr la suficiente escala de datos fue determinante ya que obstaculizó su importancia competitiva en los servicios de cartografía en el Reino Unido. Por tanto, no se puede desestimar en términos competitivos el valor de los datos. Los abusos de compañías poderosas de tecnología pueden causar grandes daños en la pérdida de la elección, la innovación, la privacidad, la autonomía y la libertad, y la confianza de los ciudadanos en una economía de mercado. Los consumidores no se benefician invariablemente cuando los servicios son gratuitos, porque no son totalmente gratuitos. Los consumidores pagan con sus datos personales y privacidad. Debido a la falta de transparencia, los consumidores a menudo no saben cuánto pagan en realidad por estos servicios.

⁵⁷ OECD Directorate for financial and Competition Committee enterprise affairs, “Big Data: bringing competition policy to the digital era,” DAF/COMP(2016)14, 2016.

⁵⁸ Para mayor información ver: CISCO White Paper: <http://www.cisco.com/c/dam/en/us/solutions/collateral/service-provider/global-cloud-index-gci/white-paper-c11-738085.pdf>

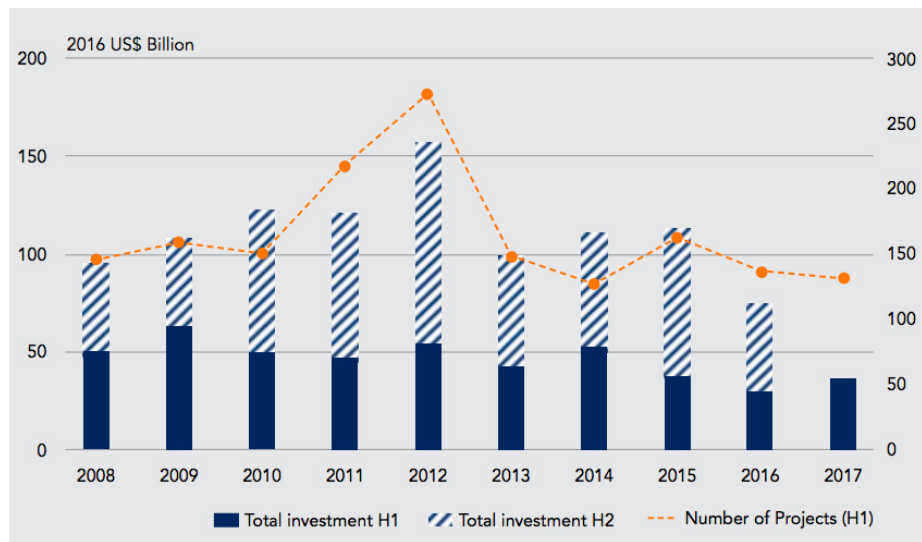
⁵⁹ La promoción de estándares abiertos en la aplicación de interfaces de programación (APIs) y formatos de datos puede ser clave para fortalecer la interoperabilidad y la reutilización de datos y servicios, mientras se fortalece la competencia de los proveedores de servicios.

⁶⁰ Private Finance Great Britain Treasury Taskforce. Partnerships for prosperity: The private finance initiative. Great Britain, Treasury, Taskforce Private Finance, 1998.

Al darle una mirada a cifras recientes, el Banco Mundial publicó el estudio de la participación del sector privado en proyectos de infraestructura (PPI, siglas en inglés) de la primera mitad de 2017. En la figura 7 se puede observar que la inversión en participación privada en infraestructura de economías emergentes y en desarrollo (EMDE, siglas en inglés) en la primera mitad del 2017 (H1) aumentó en un 24%, pasando de US \$ 29.5 mil millones en el primer semestre de 2016 a US \$ 36.7 mil millones. Es importante resaltar que en la segunda mitad del año siempre hay un mayor volumen de ejecución de recursos debido a compromisos presupuestales que se acumulan generalmente hasta final de año.

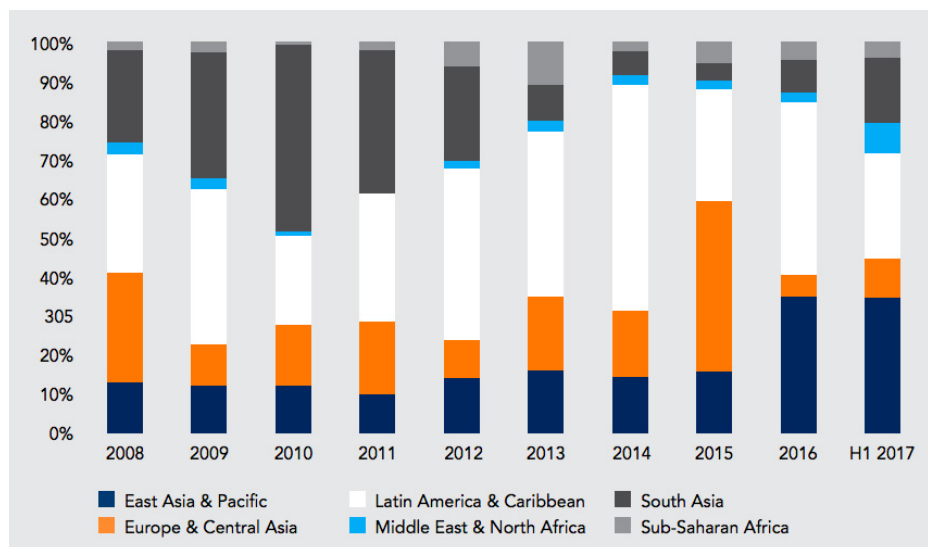
A nivel regional se puede ver que la región Latinoamericana y del Caribe tiene una de las más grandes participaciones en los últimos 10 años. Ver figura 8. Es importante resaltar que Colombia y Chile cuentan con uno de los mejores marcos para el desarrollo de APP en la región⁶³.

Figura 7. Compromisos de inversión en proyectos de infraestructura con participación privada en EMDE, 2008 H1-2017



Fuente: Enero - Junio 2017, Participación privada en infraestructura (PPI). Banco Mundial. Septiembre 2017.

Figura 8: Participación regional de los compromisos de inversión en proyectos de infraestructura con participación privada en EMDE, 2008-2016 y H1 2017



Fuente: Enero - Junio 2017, Participación privada en infraestructura (PPI). Banco Mundial. World Bank. Septiembre 2017.

Ahora bien, las APP en el desarrollo de la industria del Big Data pueden incluir entre otros los siguientes beneficios:

- Maximizar utilidad a través del uso conjunto de los recursos disponibles (información, conocimiento, presupuesto).
- Compartir los riesgos asociados al desarrollo de nuevos proyectos/negocios.
- Mejorar la competitividad a partir de la innovación.
- Aumentar el acceso a servicios e información pública (por ejemplo, a partir de Open Data).
- Contar con mecanismos de coordinación y diálogo eficientes.
- Conectar actores (nacionales con locales, empresas con gobiernos, universidades con empresas y gobiernos, etc.).
- Mejorar la calidad y la eficiencia en la entrega de servicios públicos o en procedimientos públicos administrativos. (desarrollo de negocios/plataformas/aplicaciones).
- Multiplicar el número de beneficiarios.

La cooperación, sin embargo, implica y aumenta el número de participantes. Incluso en alianzas, los actores son usualmente dependientes unos del otro. Estas dos condiciones básicas crean problemas. Un aspecto importante de las APP es la dificultad de conectarse entre los socios, dada la interacción de cada uno de los actores con sus propias redes (que involucran una serie de actores específicos ajenos a la APP), y los escenarios divergentes a los que están expuestos que hacen que su acción esté influenciada por decisiones que se toman por fuera de la alianza.⁶⁴ Los actores están demasiado preocupados por sus propios problemas y procedimientos internos, que tienen dificultades para actuar como verdaderos socios. También puede ocurrir que existan actores principales y otros que tengan un rol periférico, lo que hace que los grados de motivación y de compromiso o involucramiento sean distintos o desequilibrados. Esta situación hace que en muchos casos sea necesario establecer reglas nuevas de comportamiento para poder operar. Lo que es natural en muchas alianzas es producir situaciones en donde los esfuerzos se enfocan en separar responsabilidades y enfatizar en las relaciones contractuales. Si bien, muchas alianzas terminan siguiendo esa tendencia, la mayoría de veces, los resultados que alcanzan terminan siendo insatisfactorios para las partes.

Sin embargo, las APP también producen nuevas configuraciones, nuevas redes, creando nuevas posibilidades. Es importante tener claridad sobre los aspectos en que cada actor va a participar y la compensación que reciben en retorno a los costos que incurren. La fragmentación institucional de los proyectos APP puede crear grandes barreras que aumentan la complejidad de la toma de decisiones y llama a un esfuerzo enorme de gerencia.

Las APP que se promueven para estimular la industria Big Data deben ser entendidas como “facilitadoras/habilitadoras” (“enablers”) de iniciativas que ambos el sector público y privado buscan conseguir. Estas alianzas deberían considerar tres aspectos clave: contar con los incentivos correctos, claridad en el rol que cada actor juega, y contribuir al desarrollo de capacidades. Esto implica que cada socio debe ser franco sobre su contribución, comprometerse con hechos claros, y definir el tiempo en el cual va a estar involucrado en el proyecto. La creación de capacidad local es primordial, no sirve de nada que por ejemplo, Google cree un centro de datos en una ciudad y no haya gente que pueda mantenerlo u operarlo. Con el fin de que exista transferencia de conocimiento, los acuerdos deben ser explícitos de manera que se pueda garantizar la sostenibilidad de los proyectos.

Las alianzas que se diseñan para promover la industria Big Data, guardan muchas similitudes con las asociaciones que buscan favorecer procesos de innovación entre el sector público y privado. El gobierno debe tener claro cuáles son sus aportes y los de su/s contraparte/s, así como lo que quiere obtener de las alianzas donde está involucrado. El gobierno también debe reconocer que no es deseable promover cualquier tipo de APP dados los problemas que puede implicar la cooperación en términos de los intereses divergentes de los actores, el número de participantes, su gerencia o administración, entre otros.

Como se muestra en los siguientes casos de estudio, en una buena parte de los proyectos de Big Data, las APP tienen lugar para desarrollo de programas de formación y entrenamiento en Big Data, en otros casos, para el desarrollo de aplicaciones o proyectos de impacto social (por ejemplo, salud, transporte y seguridad)⁶⁵. Las posibilidades que se pueden derivar de estas alianzas son amplias. En algunos casos los gobiernos impulsan el desarrollo de estas alianzas desde el nivel central y de forma explícita, como Singapur lo hace, mientras que en otros casos, los gobiernos o las comunidades políticas fomentan estas alianzas en proyectos más concretos, como es el caso de la Unión Europea con la Alianza Público Privada en Big Data Value (BDVA). Algunos ejemplos de alianzas adoptadas en otros lugares del mundo se presentan a continuación.

Singapur: las Alianzas entre socios públicos y socios público-privados

En Singapur el papel que desempeña tanto el gobierno como el sector privado en la promoción de un ecosistema orientado hacia la innovación y las TIC es complementario (se refuerzan mutuamente). El papel del gobierno ha sido en esencia estimular la oferta inicial en el sector TIC en términos de la creación de oportunidades y la promoción de alianzas; tal y como

⁶¹ OECD, “Public-Private Partnerships IN PURSUIT OF RISK SHARING AND VALUE FOR MONEY”, 2008.

⁶² Ver: https://ppi.worldbank.org/~media/GIAWB/PPI/Documents/Global-Notes/PPI2017_HalfYear_Update.pdf

⁶³ Para más detalle: <https://www.fomin.org/en-us/Home/Knowledge/idPublication/190622.aspx>

⁶⁴ Los valores que persigue el gobierno y los privados pueden ser mutuamente excluyentes: los riesgos políticos en expectativas y los riesgos del mercado en cifras anuales, los procesos del sector público.

lo menciona Jeannie Chua en su artículo sobre la transformación digital de Singapur: “con un grupo dinámico, innovador y competente de empresas del sector de la información y las comunicaciones, el gobierno ha gradualmente cambiado su papel de planeador, implementador, y dueño de la información y las comunicaciones, a un rol de planificador global y facilitador, que provee un ambiente propicio y catalizador en donde las TIC pueden prosperar, crecer y difundirse. A través de alianzas innovadoras con el sector privado, el gobierno puede aprovechar la innovación del sector privado, para proteger el país de los problemas que se derivan de una tecnología obsoleta y crear empleos en el desarrollo, implementación y entrega de proyectos de TIC. Al hacerlo el gobierno es un catalizador del crecimiento de la industria.”⁶⁶

En el marco de la Estrategia digital de Smart Nation de Singapur donde están integrados todos los lineamientos para Big Data, Singapur ha establecido una red inter-organizacional de expertos en TIC en el sector público y privado. La colaboración de diferentes expertos ha contribuido a construir la capacidad tecnológica del país. Las alianzas se conforman a través de contratos formales y se centran en la colaboración permanente como un principio habilitador de proyectos concretos. Por ejemplo, la Autoridad de Desarrollo de las Tecnologías de Información y Comunicación (IDA) en todas sus iniciativas promueve siempre la colaboración de todos los actores desde el inicio de los proyectos.

El modelo de colaboración adoptado por el gobierno comprende dos formas: Alianzas entre públicos (Público-Público) y Alianzas Público-Privadas. Un ejemplo claro de alianzas entre públicos, es la existente entre la Agencia para la Ciencia, Tecnología e Investigación (A*STAR) y la Autoridad de Transporte para desarrollar un sistema de transporte público de última generación que permita mejorar la experiencia de viaje de los pasajeros. Otro ejemplo, es el Centro de Innovación Pivotal, un laboratorio creado por la Autoridad de Desarrollo de tecnologías de la información y las comunicaciones de Singapur (IDA) y la compañía Pivotal para ayudar a las empresas a desarrollar proyectos en Análisis de Big Data.

En muchos casos las entidades públicas realizan “alianzas formales con el sector privado a través de corporaciones integradas por multinacionales, empresas locales y Pymes. Un ejemplo reciente de estas APP es el Consorcio de Ciencias de Datos⁶⁷ de Singapur, promovido por la Fundación Nacional de Investigación (NRF)⁶⁸ que busca fortalecer la investigación colaborativa entre institutos y centros de investigación en la industria Big Data, para ayudar a las empresas a conocer y adoptar las últimas tecnologías.

La dinámica en la que se promueven e implementan estas alianzas está fundada en la cultura de innovación y de experimentación en la que Singapur ha construido su desarrollo social y económico. “La innovación se trata de tomar riesgos, de realizar experimentaciones cuidadosas de nuevos conceptos e ideas antes de salir en vivo. Es muy frecuente ver al gobierno de Singapur tomando el liderazgo, mientras que el sector privado y la sociedad civil ofrecen el conocimiento y realimentación, respectivamente. Este modelo de colaboración tripartita asegura que todos los actores clave en la cadena de valor estén comprometidos en el proceso de alcanzar los resultados deseados.”⁶⁹ En general, existe una apertura generalizada a la asociación, a la colaboración mutua. Esta característica particular de Singapur ha facilitado la innovación.

La Alianza Público-Privada Big Data Value (BDVA) de la Unión Europea⁷⁰

La Comisión Europea junto con la industria (grandes empresas y Pymes), los investigadores y la academia, conformaron esta alianza para cooperar en investigación e innovación en temas asociados a los Datos, fortalecer la comunidad de datos y establecer las bases de una economía orientada en Datos en Europa.

Después de firmar un acuerdo contractual en octubre de 2014, con el compromiso de una inversión de la Unión Europea de 2.5 miles de millones de euros y con la expectativa de alcanzar 2 miles de millones de euros en aportes de los socios privados, la industria (incluyendo las empresas de ATOS, Nokia Solutions and Networks, Orange, SAP, Siemens) y la academia (con centros de investigación como Fraunhofer y el Centro de Investigación Alemán para la Inteligencia Artificial) establecieron prioridades en una Agenda de Investigación e Innovación Estratégica (SRIA) a ser implementada en el período de 2016 a 2020.

Además de proyectos específicos, la APP soporta Espacios Innovadores para la experimentación con datos del sector privado y open data, que tienen previsto funcionar como incubadoras de negocios y hubs de desarrollo de capacidades y mejores prácticas. El portafolio de proyectos a desarrollar se ha dividido en las siguientes categorías⁷¹: (a) integración de datos y experimentación en proyectos multisectoriales, (b) acciones piloto de gran escala en los sectores que más se benefician de la innovación orientada en datos (como bio data y transporte), tecnologías para la protección de la privacidad de datos, investigación e innovación responsable en TIC, soporte, desarrollo de capacidades industriales, benchmarking y evaluación⁷².

⁶⁵ Con referencia a los derechos de propiedad en APPs de Big Data, hay varias formas de manejarlas, la primera son las licencias de propiedad intelectual, la segunda son las patentes comunes y plataformas de licencias y la licencia abierta.

⁶⁶ Hanna y Knight, National strategies to harness information technology: Seeking transformation in Singapore, Finland, the Philippines, and South Africa. Springer Science + Business Media. 2012.

⁶⁷ Ver: <http://sdsc.sg/>

⁶⁸ La APP está conformada por la Universidad Nacional de Singapur (NUS), La Universidad Tecnológica de Nanyang (NTU), la Universidad de Negocios de Singapur (SMU) y la Agencia de Ciencia, Tecnología e Investigación (A*STAR)

⁶⁹ Ver a: Liang Foo y Gary Pan. Singapore's vision of a smart nation. 2016.

⁷⁰ Para mayor información sobre su justificación: http://europa.eu/rapid/press-release_MEMO-14-583_en.htm

Características principales de la Alianza Público-Privada Big Data Value (BDVA) de la Unión Europea

¿Para qué se necesita una APP para datos? Actualmente, apenas 2 de las 20 compañías más importantes del mundo que cambian vidas y ganan dinero de Big Data son europeas. Europa necesita invertir y fortalecer todas las partes de la “cadena de valor de datos”.

¿De qué trata la APP? Es una alianza entre la Comisión Europea y la BDVA y conecta la industria europea, investigadores, académicos, y la Comisión Europea a fin de cooperar en la investigación e innovación de datos. La Comisión responderá a los retos y necesidades principales identificados en SRIA. Se espera que las inversiones lleguen a €2.5 mil millones durante 2016-2020 y la Comisión ha destinado un presupuesto de aproximadamente €500 millones, de los cuales se espera que cada euro genere 4 euros de inversión del sector privado (alrededor de 2 000 millones de euros).

¿Cómo puede la APP beneficiar las industrias tradicionales o en gran escala y las Pymes? La APP incluirá líderes de varios sectores de industria europea y ayudará a identificar procesos industriales de gran volumen y alto valor que hoy son ineficientes debido a la falta de acuerdo sobre los formatos y procesos de datos. Además, el objetivo de la APP es que al menos el 20% de las organizaciones que participan en sus proyectos sean Pymes. La APP facilitará a las Pymes acceder a grandes conjuntos de datos de grandes empresas y organizaciones, lo que puede ser beneficioso para todos.

¿Cuál es su impacto en los trabajos e innovación? Dará lugar a la creación de cientos de miles de trabajos nuevos en Europa. Según un estudio⁷³, en Reino Unido, se pronostica que el número de personal especialista en Big Data aumentará en un 243% durante los próximos 5 años a aproximadamente 69.000 personas. Los “espacios de innovación” (i-Spaces) servirán de incubadoras de negocios y centros para el desarrollo de capacidades, y les darán acceso a los estudiantes de las universidades e instituciones de investigación acceso a conjuntos de datos grandes y realistas y herramientas de análisis de datos de código abierto.

¿Cómo se relaciona con la seguridad y la privacidad? La Comisión seguirá las prioridades establecidas por la “Agenda Estratégica de Investigación e Innovación” de la industria Big Data y financiará nuevas investigaciones sobre soluciones técnicas para integrar características de mejora de la seguridad y la privacidad “por diseño”. La Comisión también proporcionará orientación sobre cómo las empresas pueden hacer anónimos los datos y utilizar seudónimos para realizar análisis de riesgos de datos personales, y sobre las herramientas e iniciativas disponibles para mejorar la conciencia del consumidor.

¿Cómo abordará la APP las barreras y preocupaciones legales relacionadas con los datos?

Características principales de la Alianza Público-Privada Big Data Value (BDVA) de la Unión Europea

Lo que es importante resaltar de esta alianza BDVA es que ha logrado integrar a un número elevado de actores de diferente naturaleza y sector. Estos actores a su vez se han organizado para poder participar, creando por ejemplo, en el caso de las empresas, una asociación autónoma que las representa ante los demás socios de la Alianza (la Big Data Value Association). Otro aspecto relevante, es que esta alianza hace parte del Programa Horizonte 2020 de la Unión Europea que tiene un presupuesto de 80 miles de millones de euros y que además de otros instrumentos, usa y promueve las alianzas y la colaboración entre diferentes actores, para la realización de proyectos específicos. En la mayoría de los casos, para que un proyecto sea aprobado, la constitución de alianzas es un requisito explícito.

Los recursos importantes con los que cuenta la Alianza son sin duda uno de los atractivos más importantes para los socios, y por tanto su implementación requiere contar con unos lineamientos claros y una gobernabilidad eficiente. Si bien, hacer realidad esta Alianza en el marco de La Unión Europea, ha implicado una serie de tiempos y procedimientos administrativos largos, situación que ha hecho que su implementación apenas haya iniciado en 2017, esta Alianza hoy cuenta con una estructura organizacional y operacional definida, que le permite ser administrada y evaluada. Este es un aspecto esencial de las APP como instrumentos efectivos de colaboración. Aquí son los consorcios mismos que deciden y no hay un rol tan fuerte de los gobiernos, como ocurre en el caso de países como Singapur que participan activamente en todo los procesos de asociación.

⁷¹ Para mayor información sobre estos proyectos ver: <http://www.bdva.eu/?q=ppp-projects>

⁷² http://europa.eu/rapid/press-release_IP-14-1129_en.htm

⁷³ E-Skills UK. “Big Data Analytics An assessment of demand for labour and skills, 2012-2017”. Enero 2013.

⁷⁴ Ver: <https://www.bigdata-alliance.org/about/>

⁷⁵ Esta alianza está conformada por un consorcio de organizaciones privadas, que incluyen MedCom e IBM junto con las autoridades danesas del sistema de salud. DIGILE es una compañía sin ánimo de lucro que busca aumentar el conocimiento y desarrollo de herramientas requeridas por la economía de Internet (creando crecimiento, trabajos y nuevos negocios en Finlandia) a partir de programas de investigación entre actores del sector público y privado. Para mayor información sobre Digile ver: <http://www.internetofthings.fi/>

Otras Alianzas

En otros países las APP toman enfoques globales y/o sectoriales. Por ejemplo, en Países Bajos, la Plataforma Big Data Alliance⁷⁴ fue creada para promover la colaboración entre instituciones de la academia y la industria de Big Data y Análisis de Datos. Fue fundada por: Universidad de Amsterdam, CWI (instituto de investigación en matemáticas e informática), SURFsara, y ORTEC. Se enfoca en el público en general, estudiantes, negocios e investigadores. El objetivo consistía en contar con un mecanismo para compartir conocimiento, experiencia e investigación y estimular la innovación y la educación en el campo de Big Data. Reuniendo actores clave de la academia y la industria en una misma plataforma, se facilita la colaboración mutua y el progreso en tanto la investigación como en los negocios en Big Data. En varios países de Europa existen al mismo tiempo un sinnúmero de ejemplos de alianzas de tipo sectorial o temática, algunos de ellos son: BIT4health (better IT for better health) en Alemania, Portal e-health (sundhed.dk) en Dinamarca, DIGILE Internet of Things⁷⁵ en Finlandia.

Otro modelo y ejemplo de una asociación público-privada es el proyecto OPAL - el primer intento de gran escala de este tipo - que ya se mencionó en la primera parte de este documento - está descrito en más detalle en el cuadro a continuación. OPAL es considerable en varios aspectos;

En primer lugar, como se discutió en la sección 1, proporciona los medios necesarios para convertir los ‘datos masivos’ en ‘datos abiertos’. Segundo, e igualmente fundamental, su objetivo es cerrar las brechas y fomentar conexiones, capacidades y confianza entre las comunidades de Big Data y Open Data - y el ecosistema de datos en toda su extensión. Hay varias maneras en que hace esto, como se describe en la tabla a continuación.

Promover el uso de los datos, la confianza, y las capacidades a través de plataformas seguras: OPAL

El proyecto Open Algorithms (OPAL) es una innovación socio-tecnológica desarrollada por un consorcio compuesto del MIT Media Lab, Data-Pop Alliance, World Economic Forum, Orange, Imperial College London y Overseas Development Institute, para aprovechar los datos provenientes del sector privado para el bien común por “enviar el código a los datos” de una manera que preserve la privacidad, sea predecible, participativa, escalable y sostenible.

Esta plataforma está diseñada para proporcionar una imagen más precisa y matizada de la realidad humana a los estadísticos oficiales, políticos, planificadores, empresas y ciudadanos mientras que permite una mayor inclusión y aportes de los individuos, con respeto a los tipos y el uso del análisis realizado sobre sus datos. OPAL se basa en años de trabajo en el MIT y en otros lugares, y es un hito clave hacia la realización de una visión donde los datos están en el centro del desarrollo social en todo el mundo y apoyan los Objetivos de Desarrollo Sostenible de las Naciones Unidas y la democracia.

OPAL surgió del reconocimiento de que el acceso a las fuentes de Big Data en manos de las empresas privadas con fines de investigación y política ha sido un acertijo. Hasta la fecha, se ha accedido y analizado externamente uno de los más valiosos tipo de datos (conocidos como Call Detail Records (CDR)) recopilados por operadores de telecomunicaciones con fines de facturación a través de ‘desafíos de datos’ como Desafíos D4D de Orange o acuerdos bilaterales en virtud de Acuerdos de No Divulgación. Compromisos de este tipo han ofrecido una amplia evidencia de su potencial y su demanda. Una gran cantidad de literatura académica ha mostrado cómo el análisis computacional de los CDR (tanto como los datos bancarios), junto con datos de encuestas tradicionales y otras estadísticas oficiales, pueden ayudar a capturar los resultados y procesos socioeconómicos en altos niveles de granularidades geográficas y temporales y grados de complejidad - incluida la propagación de enfermedades, la pobreza, la alfabetización, la delincuencia, así como la optimización de la prestación de servicios públicos y los sistemas de transporte, en particular.

Al mismo tiempo, existen riesgos asociados con el uso de datos tan personales y conectados. Las preocupaciones por la privacidad y la seguridad han crecido a medida que la noción de anonimización de datos fue probado cada vez más, como lo demuestra la investigación del MIT, y la naturaleza y el alcance de las actividades de vigilancia de la Agencia de Seguridad Nacional (NSA) que Edward Snowden reveló públicamente, alimentan temores de un futuro Orwelliano. Las posibilidades de desequilibrios crecientes entre los grupos que tienen acceso a los datos y las capacidades y los que no, y la consiguiente concentración de poder, también son preocupantes. Una crítica relacionada es que los algoritmos que se usan cada vez más para tomar decisiones de política son similares a las “cajas negras” que ocultan reglas y procedimientos que no pueden someterse al escrutinio y reparación pública.

A la luz de estos obstáculos y requisitos, la visión de OPAL es el desarrollo de un nuevo tipo de sistema tecnoinstitucional basado en la confianza que aproveche de datos del sector privado para fomentar la transparencia, la agilidad, la rendición de cuentas y la inclusión respetando la privacidad y la seguridad. Con ese fin, OPAL reflejará y fomentará un cambio paradigmático para dar la vuelta a Big Data completamente y salvarlo de sí mismo. El primer paso es enviar los algoritmos a los datos, no al revés, para que los datos no estén expuestos a robos y mal uso. El segundo paso es el diseño inclusivo de cómo se usan los algoritmos de big data, para que sirvan a las necesidades locales y respeten los estándares locales, en lugar de imponer perspectivas y experiencias externas.

OPAL está comenzando con 2 pilotos en Senegal y Colombia financiados por la Agencia Francesa de Desarrollo (AFD), en asociación con sus respectivas Oficinas Nacionales de Estadística (ANSD y DANE), así como el Departamento de Planeación Nacional en Colombia (DNP) y las principales telecomunicaciones locales operadores-Orange-Sonatel y Telefónica Colombia. En la fase de implementación inicial de OPAL, las solicitudes de indicadores predeterminados, por ejemplo, densidades de población, se pasarán por la plataforma a través de algoritmos pre-desarrollados que se ejecutan en los servidores de datos de las compañías, detrás de sus firewalls, de una manera que preserve la privacidad, y los resultados estarán disponibles a través de una interfaz. La participación y el empoderamiento local serán fundamentales para el desarrollo de OPAL: las necesidades, los comentarios y las prioridades se recogerán e identificarán mediante talleres y debates locales, y sus resultados se incorporarán al diseño de futuros algoritmos. Estos algoritmos estarán abiertos, por lo tanto, estarán sujetos al escrutinio y reparación pública. Una junta asesora local proporcionará orientación y supervisión al proyecto, para garantizar que se atenga a los principios éticos clave. Además, se impartirá capacitación en torno al proyecto para fomentar su uso y difusión, así como las capacidades y conexiones de manera más amplia. Si los dos pilotos iniciales se consideran exitosos, el proyecto podría extenderse a otros países en desarrollo. Recientemente, en junio de 2017, se llevó a cabo en Bogotá el primer Taller Inaugural de OPAL en Colombia, en la que reunieron 35 profesionales de diferentes sectores. El evento contó con representantes del Gobierno nacional, el Gobierno distrital, la sociedad civil, el sector privado, la academia, Pymes y organizaciones internacionales.

Fuente: Alex Pentland y Emmanuel Letouzé

Las alianzas que se promuevan en muchos países para fortalecer la industria consideran en la mayoría de los casos todos los sectores y no sólo los del sector TIC o High Tech, partiendo de la premisa de que es necesario involucrar a todos los sectores y a las empresas de todos los tamaños. En muchas alianzas, los gobiernos son los iniciadores, moderadores o facilitadores. Por ejemplo, el gobierno puede iniciar un diálogo entre múltiples actores, tomar una acción colectiva o desarrollar esfuerzos de construcción de capacidad con las empresas, involucrar diferentes socios en procedimientos de estandarización normativa o simplemente movilizar recursos financieros. Un gran número de alianzas se han establecido en las últimas décadas para solucionar problemas de reducción de la pobreza, acceso a la salud, la educación, entre otros.” Otras alianzas que los países pueden considerar son las que están orientadas a integrarse a procesos de colaboración internacional, como el caso de Open 500, que cada vez aumenta su red de países y asociados en torno a Big Data.

3.2. FACILITAR EL DESARROLLO DE CLUSTERS A NIVEL LOCAL

El desarrollo de una empresa está determinado en buena medida por su entorno: las instituciones, la infraestructura y la agrupación de otras empresas a su alrededor son factores que influyen directamente en su desempeño. La productividad y la innovación que pueden alcanzar las empresas difieren si estas se encuentran en un entorno donde existe acceso a servicios, programas académicos, proveedores, infraestructura, entre otros. Contar con instituciones fuertes y reglas claras para asegurar mercados abiertos y transparentes influye positivamente en el desempeño de las empresas.

El desarrollo de clusters para el sector TIC puede dar algunas pistas sobre las posibilidades que existen para la industria Big Data. Se trata de un círculo virtuoso, donde se eliminan o mitigan las deficiencias que existen en las condiciones externas para las empresas (que se convierten en costos internos para éstas) y se generan oportunidades de crecimiento y mejora de tipo económico y social. Las políticas que se han mencionado a lo largo de este documento, como incentivos fiscales, créditos subsidiados, inversión en infraestructura, regulación y programas especiales para start-ups pueden ayudar a impulsar la industria Big Data en lugares que en muchos casos no son el referente de desarrollo tradicional.

3.2.1 FACILITAR EL DESARROLLO DE “BIG DATA HUBS”

Pensar en clusters sostenibles para la industria Big Data implica asegurar mecanismos que faciliten la innovación (por ejemplo medidas de incentivo fiscal o económico) y contar con un capital humano que pueda ser proveído (por ejemplo por universidades en los lugares donde están establecidas las empresas). La provincia de Guizhou en China es un ejemplo de cómo es posible transformar una provincia atrasada (“el PIB per Cápita de Guizhou en 2014 era de USD 4300 y a 2017 de USD 4987”, hoy en día oscila) en un hub para Big Data.

Por un lado, se pusieron en práctica medidas que atraieron empresas innovadoras a Guizhou, en particular reducciones en impuestos y subvenciones a las firmas de tecnología. “Las empresas Big Data que cumplían ciertos requisitos estaban exentas del impuesto sobre la renta (las ganancias) los tres primeros años y se beneficiaban del 50 por ciento de reducción en ese mismo impuesto los siguientes tres años.” Por otro, también se ofrecieron “subsidios de vivienda para atraer talento Big Data.”⁷⁶

Aunque esta provincia de China con una población de 39 millones de habitantes y del tamaño de Uruguay está haciendo una apuesta importante, al desarrollar una zona industrial de alta tecnología y con el establecimiento del primer laboratorio de ingeniería en Big Data del país, sólo la implementación de medidas en el mediano y largo plazo, van a permitir a Guizhou

alcanzar niveles de desarrollo sostenibles. Guizhou cuenta hoy con universidades establecidas, pero sin experiencia en el sector TIC. Estos son algunos aspectos que tanto la provincia de Guizhou en China como otras regiones del mundo que buscan convertirse en verdaderos clusters de Big Data deben contemplar.

3.2.2 EMPODERAR A LOS CIUDADANOS PARA CONSOLIDAR UN ECOSISTEMA DE BIG DATA SEGURO Y JUSTO

Los gobiernos necesitan promover el uso responsable de los datos personales a través de la promoción de tecnologías que ayuden a los individuos a apoderarse de sus propios datos, asegurando al mismo tiempo su privacidad. Existen algunos ejemplos en diferentes lugares del mundo que buscan lograr el apoderamiento de los individuos sobre sus propios datos: Self Data en Francia, Mydata en Finlandia y Midata en Reino Unido (también Vendor RelationShip Management-VRM, Cloud Personnel, Quantified Self, Personal Data Store (PDS), Personal Information Management System- PIMS, Customers Commons).

- **MIDATA en UK:** Un programa lanzado en abril de 2011 dentro de la Estrategia de Empoderamiento del Consumidor del Gobierno. Incluye empresas, grupos de consumidores y el gobierno. La idea es que los individuos puedan ver, acceder y usar sus datos personales de consumo y de transacciones de una manera que sea fácil y segura. Esto ayuda a los consumidores a sacar provecho de aplicaciones y servicios en los cuales ellos pueden usar esos datos para encontrar mejores precios o entender sus hábitos de gasto.

A través de un portal virtual se provee información pública relacionada con la industria líder en el esquema en el cual los proveedores participan (bancos o sociedades inmobiliarias o de construcción proveen cuentas personales corrientes o PCA que le permiten a los consumidores descargar sus datos de transacciones en un archivo midata. Este archivo es un registro de la historia de las transacciones de los consumidores bajo estándares establecidos. Una vez el consumidor ha descargado sus datos, estos pueden ser enviados a comparadores de precio (páginas web donde se pueden comparar precios). Estos comparadores analizarán los datos en un archivo y proveerán información a la medida para el consumidor.

- **MES INFOS en Francia:** Una persona obtiene sus datos personales de las organizaciones con los que se encuentra en relación y adiciona los datos que esa persona captura o produce por sí mismo (actualmente estos datos están en su mayoría en manos de empresas y son difíciles de obtener por los individuos). Luego la plataforma los guarda y los administra de manera segura (por ejemplo a través de *Pims Cloud Personnels*). Finalmente, el individuo explota sus propios datos con la ayuda de aplicaciones y de servicios para conocer mejor su comportamiento, tomar las mejores decisiones, evaluar sus decisiones pasadas, etc.
- **My Data en Finlandia:** La idea central es permitir a los individuos tener el control de sus propios datos. Esto simplifica el flujo de los datos y abre nuevas oportunidades para que los negocios desarrollen servicios innovadores basados en datos personales, preservando al mismo tiempo la privacidad. También han desarrollado una Red Global de MyData que busca estructurar y empoderar un movimiento global a través de reuniones, conferencias y otras interacciones con nuevos Hubs Globales que se vayan creando alrededor en torno al empoderamiento de los individuos sobre sus datos.

3.2.3 LANZAR INICIATIVAS PÚBLICAS QUE CONECTEN A LOS ACTORES DEL ECOSISTEMA

Además del diseño y la implementación de programas nacionales a gran escala, como se describió anteriormente, un gobierno también puede lanzar iniciativas que conecten directamente grupos de usuarios, organizaciones o empresas. La expansión de redes comerciales (geográficamente y por escala); la facilitación de la comunicación y el trabajo entre las empresas, los proveedores (de bienes o servicios), los talentos y los ciudadanos; el desarrollo de la información disponible en el mercado (nuevas tecnologías, nuevas áreas de conocimiento); entre otros representan enfoques y perspectivas impactantes que pueden llevar a la innovación y el crecimiento orgánico.

Singapur es un buen ejemplo del éxito de este modelo; ya ha llevado a cabo el desarrollo de iniciativas específicas en Big Data, hackatones, plataformas y redes tecnológicas, incluso incentivos para que empresas emergentes creen aplicaciones que los ciudadanos puedan utilizar con datos extraídos del gobierno (en sectores como el del transporte y la salud).⁷⁷ Algunos ejemplos de iniciativas y proyectos en el sector TI y donde subyacen proyectos Big Data son: SG Innovate, One-North, Jurong Innovation District (JID) y Build Amazing Start-ups Here (BASH). Desde parques de negocios e incubadoras hasta entorno de pruebas (sandbox), hoy, Singapur desarrolla proyectos ambiciosos que abarcan grandes inversiones para los sectores público y privado. Esta ola de iniciativas es parte de la misión más amplia del gobierno de Singapur, como se describe en mayor detalle en el Documento 1, para establecer la ciudad-estado como un centro líder de desarrollo e innovación de Big

⁷⁶ Nir Kshetri, Torbjorn Fredriksson, y Diana Carolina Rojas Torres, *Big Data and Cloud Computing for Development: Lessons from Key Industries and Economies in the Global South*. Taylor & Francis, 2017.

⁷⁷ Por ejemplo, SG Innovate fue creado para incubar empresas emergentes vinculadas a la ciencia y la tecnología al conectar a empresarios con mentores y permitirles el acceso a institutos de investigación, presentarles las oportunidades de financiamiento y eventualmente exponerlos a los nuevos mercados. Los inversionistas del sector privado, las compañías tecnológicas y las instituciones académicas son alentadas activamente por el gobierno para participar en la iniciativa financiera o técnicamente.

OBSERVACIONES

Los gobiernos juegan un rol clave en el desarrollo del ecosistema y en la industria de Big Data, teniendo como tarea principal crear un contexto tecno-político o socio-tecnológico que facilite el fortalecimiento del ecosistema compuesto por las 3 C de Big Data (“Crumbs” o Migajas, Capacidades, y Comunidad/Cultura).

Por un lado, el gobierno tiene la responsabilidad de garantizar el bien público y la justicia. Los abusos de compañías poderosas de tecnología pueden causar grandes daños en la pérdida de la elección, la innovación, la privacidad, la autonomía y la libertad, y la confianza de los ciudadanos en una economía de mercado. Asimismo, cabe al estado promover prácticas éticas para el uso de Big Data y proponer marcos de intercambio de datos a la vez seguros, innovadores y productivos.

Por otro lado, como vimos a lo largo de este documento, el estímulo de la industria y la promoción de la innovación crea valor económico y potencialmente social. Exploramos diversas medidas que puede tomar el gobierno para estimular las capacidades a lo largo de toda la vida de los ciudadanos. La perspectiva presentada en este documento es la de una de alta colaboración entre el sector privado y las políticas gubernamentales. Se trata de desarrollar una *empatía profunda* con las empresas para conocer sus problemas, su lógica de funcionamiento y sus redes, para así generar sinergias entre la demanda y oferta de capacidades. Algunas iniciativas que fomentan dicha sinergia pueden incluir incubación de nuevas empresas, colaboración entre empresas y universidades a través de centros de creación de conocimiento y zonas de innovación, promoción de una fuerza de trabajo a través de aprendices y otros mecanismos entre la industria y la educación.

Por fin, el gobierno tiene un rol clave en la configuración de un ecosistema de Big Data saludable, incluyendo la creación de sinergias entre los actores involucrados, desde ciudadanos hasta generadores de datos, analistas y usuarios finales, pasando por las instituciones socio-políticas (leyes, cuadros político-institucionales). Asimismo, concebimos un gobierno puente que conecta y que permite a los otros actores realizar su papel dentro del ecosistema. Las posibilidades que un gobierno puente tiene para estimular la industria Big Data son muy amplias, algunos ejemplos incluyen conectar pequeñas y grandes empresas para promover el crecimiento y éxito de pequeñas y medianas empresas; revitalizar grandes compañías a través de alianzas con Pymes innovadoras; conectar líderes en todos los sectores para desarrollar estrategias regionales y producir modelos escalables que se construyan en las bases locales y puedan atraer nuevas inversiones.

El objetivo de este documento fue describir y analizar por qué, cuándo, cómo y con qué resultados los gobiernos de los países seleccionados implementaron medidas para estimular el sector naciente de Big Data, con el objetivo de generar impacto en la industria y en la economía del país. Con base en el análisis de la situación colombiana que será presentada en el Documento 3, los elementos considerados relevantes serán profundizados y aterrizados para la formulación de la estrategia nacional en el Documento 4.

BIBLIOGRAFÍA

- Acquisiti, Alessandro. "The Economics of Personal Data and the Economics of Privacy". *Carnegie Mellon University, Research Showcase @ CMU*, 2010.
- Aggarwal, Devi Dayal. *History and development of elementary education in India*. Vol. 1. Sarup & Sons, 2002.
- Averou, Chrisanthi. "The link between ICT and economic growth in the discourse of development". London School of Economics.
- Banco de Desarrollo de América Latina, 2013, Technology OECD Directorate for Science and Innovation, "Data-driven innovation for growth and well-being," *OECD Interim Synthesis Report*, 2014.
- Banco de Desarrollo de América Latina, *Hacia la transformación digital de América Latina: las infraestructuras y los servicios TIC en la región*. Banco de Desarrollo de América Latina, 2013.
- Banco de Desarrollo de América Latina. "Hacia la transformación digital de América Latina: las infraestructuras y los servicios TIC en la región." *Banco de Desarrollo de América Latina*, 2013.
- Big Data Value Association. "BDV Factsheet". Disponible : bdva.eu/sites/default/files/141010_factsheet_bdv_cppp_final.pdf
- Commission, Indian Education, et al. "Government of India, Ministry of Human Resource Development, educational statistics at a glance." *New Delhi*, 2016.
- Conner-Simons, Adam. "System predicts 85 percent of cyber-attacks using input from human experts", MIT News. April 18, 2016.
- Conrad, Larry y Rivers, Paul. "Protect Yourself: Staying Safe in Cyberspace", Technology @ Berkeley. October 6, 2015. Disponible: <https://technology.berkeley.edu/news/protect-yourself-staying-safe-cyberspace>
- Deloitte and ASSOCHAM India. "Digital India: Unlocking the trillion dollar opportunity," 2016.
- Dini, Marco, Giovanni Stumpo, and Italia Cooperazione Italiana. "Políticas para la innovación en las pequeñas y medianas empresas en América Latina," 2011.
- Economist, The. "the future of Indian business education." *The Economist*, 2016.
- financial, OECD Directorate for, and Competition Committee enterprise affairs. "Big Data: bringing competition policy to the digital era." *DAF/COMP(2016)14*, 2016.
- FOO, Liang, y Gary PAN. "Singapore's vision of a smart nation," 2016.
- Future of Privacy Forum, Understanding corporate data sharing decisions: practices, challenges, and opportunities for sharing corporate data with researchers, Noviembre 2017
- Gollakota Kamala, Pick, James B and Sathyapriya. Using technology to alleviate poverty: use and acceptance of telecenters in rural India. Volume 18. Issue 3. 2012.
- Great Britain Treasury Taskforce, Private Finance. *Partnerships for prosperity: The private finance initiative*. Great Britain, Treasury, Taskforce Private Finance, 1998.
- Hanna, Nagy K, and Peter T Knight. *National strategies to harness information technology: Seeking transformation in Singapore, Finland, the Philippines, and South Africa*. Springer, 2011.
- Information Commissioner's Office (ICO). "Big data, artificial intelligence, machine learning and data protection", Data Protection Act and General Data Protection Regulation. Versión 2.2, 2017.
- International Data Corporation (IDC). *Double-Digit Growth Forecast for the Worldwide Big Data and Business Analytics Market Through 2020 Led by Banking and Manufacturing Investments, According to IDC*, October 2016. <http://www.idc.com/getdoc.jsp?containerId=prUS41826116>.
- International Labour Organisation (ILO). "A skilled workforce for strong, sustainable and balanced growth: a G20 training strategy," 2010.
- Kanter, Rosabeth Moss. "Enriching the ecosystem." *Harvard Business Review* 90, no. 3 (2012): 140-+.
- Klijin, Erik-Hans, and Geert R Teisman. "Institutional and strategic barriers to public—private partnership: An analysis of Dutch cases." *Public money and Management* 23, no. 3 (2003): 137-146.

- Kshetri, Nir, Torbjorn Fredriksson, and Diana Carolina Rojas Torres. *Big Data and Cloud Computing for Development: Lessons from Key Industries and Economies in the Global South*. Taylor & Francis, 2017.
- Linden, Leigh, *Complement or Substitute? The Effect of Technology on Student Achievement in India*, Columbia University, 2008.
- Manyika, et al. McKinsey Global Institute. "Big data: The next frontier for innovation, competition, and productivity." 2011.
- Mateos-Garcia, J, G Windsor, and S Roseveare. "Analytic Britain: Securing the right skills for the data-driven economy." *London: Nesta*, 2015.
- Mazzarol, Tim. *6 ways governments can encourage entrepreneurship | World Economic Forum*, December 2014. Disponible: <https://www.weforum.org/agenda/2014/12/6-ways-governments-can-encourage-entrepreneurship/>.
- Obijiofor, Levi. *New Technologies in Developing Societies. From Theory to Practice*. Palgrave Studies in Communication for Social Change. 2015.
- OECD(2017),OECDDigitalEconomyOutlook2017,OECDPublishing,Paris.Disponible:<http://dx.doi.org/10.1787/9789264276284-en>
- OECD Directorate for financial and Competition Committee enterprise affairs, "Big Data: bringing competition policy to the digital era," *DAF/COMP(2016)14*, 2016.
- OECD Directorate for Science, Technology, and Innovation. "Data-driven innovation for growth and well-being." OECD INTERIM SYNTHESIS REPORT, 2014.
- OECD, "Public-Private Partnerships In Pursuit Of Risk Sharing And Value For Money", 2008.
- ONeil, C. (2016). *Weapons of math destruction: how big data increases inequality and threatens democracy*. New York: Crown.
- Overy, Allen. "The EU General Data Protection Regulation," 2017.
- Plomp, T.J., R.Anderson, N. Law and A. Quale. International Association for the Evaluation of Educational Achievement (IEA) *Cross national information and communication technologies policies and Practices in Education*. Chapter 19.P.3. Formation Age Publishing- IAP.
- Ponemon Institute. 2016 Cost of Cyber Crime Study & the Risk of Business Innovation, Benchmark Study of 237 Global Companies Ponemon Institute October 2016. Disponible: <https://software.microfocus.com/en-us/asset/2016-cost-cyber-crime-study-risk-business-innovation-ponemon-institute-cyber-security-analysis>
- Private Finance Great Britain Treasury Taskforce. *Partnerships for prosperity: The private finance initiative*. Great Britain, Treasury, Taskforce Private Finance, 1998.
- Schanzenbach, Diane. *America's small investment in government data has big payoffs*, March 2017. <https://www.brookings.edu/opinions/americas-small-investment-in-government- data-has-big-payoffs/>.
- Singhi, Namrata. "Data Scientists Earning more than CAs, engineers." *The Times of India*, 2016.
- Union, European. "Regulation (EC) No 679/2016 of the European Parliament and of the Council of the 27 April 2016 The General Data Protection Regulation." OJ L 119 (2016): 33–34.
- Wachter, Mittelstadt, Floridi. "Why a right to explanation of automated decision-making does not exist in the General Data Protection Regulation", *International Data Privacy Law*, 2017. University of Oxford, Oxford Internet Institute. (2016)
- World Bank, World Development Indicators. Disponible: data.worldbank.org/data-catalog/world-development-indicators
- Zarsky, Tal, *Incompatible: The GDPR in the Age of Big Data* (August 8, 2017). *Seton Hall Law Review*, Vol. 47, No. 4(2), 2017. Disponible: <https://ssrn.com/abstract=3022646>

