

ESTUDIO DEL GRADO DE PREPARACIÓN PARA LA APERTURA DE DATOS DEL SECTOR EDUCATIVO EN COLOMBIA



NOVIEMBRE DE 2015



vive digital
para la gente





ÍNDICE

INTRODUCCIÓN		5
RESUMEN EJECUTIVO		6
METODOLOGÍA		13
A. METODOLOGÍA PARA APOYAR INICIATIVAS SECTORIALES DE DATOS ABIERTOS		17
A.1. ¿Por qué realizar un análisis sectorial?	13	
A.2. Diagnóstico estratégico como punto de partida	13	
A.3. Análisis del sector	14	
A.4. Informe final	14	
B. LA EDUCACIÓN EN COLOMBIA: SITUACIÓN DE PARTIDA Y METAS A ALCANZAR		17
B.1. Situación de Colombia respecto de los países de la OCDE	17	
B.2. Clasificación de los grandes retos	20	
B.3. Metas de Colombia para el sector educación	21	
C. BARRERAS Y OPORTUNIDADES EN EL SECTOR EDUCATIVO		25
C.1. Trabajo de campo	25	
C.2. Barreras para la política de apertura y reutilización de datos	26	
C.3. Oportunidades para la política de apertura y reutilización de datos	32	
D. BENEFICIOS DE UNA POLÍTICA DE APERTURA Y REUTILIZACIÓN DE DATOS		37
D.1. Para qué retos pueden ser útiles los datos abiertos en Colombia	37	
D.2. Posibles beneficios para Colombia detectados en el estudio de campo	39	
E. ANÁLISIS DE FUENTES DE DATOS PUBLICABLES		45
E.1. Categorización de las fuentes de datos	45	
E.2. Publicación actual de datasets en el portal de datos abiertos de Colombia	46	
E.3. Comparativo internacional de publicación de datos abiertos del ámbito educativo	47	
E.4. Conjuntos de datos potencialmente publicables como Open Data	50	
F. ESTUDIO DE LA DEMANDA		55
F.1. Demanda interna de datos de educación	55	
F.2. Demanda externa de datos de educación	60	
G. RECOMENDACIONES PARA LA ACCIÓN		62
G.1. Recomendaciones generales	63	
G.1. Recomendaciones sobre datos	64	
ANEXOS		65
ANEXO 1: Personas e instituciones entrevistadas	66	



.INTRODUCCIÓN

Este informe ha sido preparado por el Banco Mundial en el marco del convenio de asistencia técnica encargado por el Ministerio de Tecnologías de la Información y las Comunicaciones (MinTIC) del Gobierno de la República de Colombia, cuyo objetivo se centra en el fortalecimiento del ecosistema de innovación digital en Colombia.

El equipo de proyecto del Banco Mundial estuvo liderado por Eva Clemente, Arturo Munte Kunigami, Henry Forero, y compuesto por Rocio Sánchez Viguera, Mónica Peñuela, Claudia Ivette García Romero, Jean Barroca, Samhir Vasdev, Nagore de los Ríos, Alberto Ortiz de Zárate, Carlos de la Fuente, Martha Peña, Hallie R. Applebaum, Mikko Korja, y Fabio Andrés Montoya.

Este informe ha sido elaborado por los expertos en Datos Abiertos y Gobierno Abierto del Banco Mundial, Carlos de la Fuente, Alberto Ortiz de Zárate y Nagore de los Ríos. Este informe no hubiese sido posible sin el apoyo del equipo encargado en el Gobierno de Colombia del impulso de la Iniciativa de datos abiertos, Diego Barbosa, Liliana Jaimés y Luisa Fernanda Medina

dirigido por Johanna Pimiento, directora de Gobierno en Línea del Viceministerio de TI liderado por María Isabel Mejía del Ministerio de Tecnologías de la Información y las Comunicaciones de la República de Colombia. Igualmente hay que mencionar la estrecha colaboración de numerosas personas del MinEducación (MinEducación), bajo la coordinación de Víctor Saavedra y Carlos Lugo. También hay que resaltar la colaboración del Instituto Colombiano para la Calidad de la Educación y de otras entidades del sector educación.

Este estudio incluyó varias entrevistas a profundidad con representantes de gobierno y sociedad civil organizada, durante los meses de octubre y noviembre de 2015. Se ha incluido una lista de las personas e instituciones entrevistadas como anexo del informe.

Las conclusiones e interpretaciones expresadas en esta publicación no reflejan necesariamente las opiniones del Banco Mundial o los gobiernos que representa el Banco Mundial.

.RESUMEN EJECUTIVO

OBJETIVO

El objetivo de este estudio es conocer la situación actual de los conjuntos de datos vinculados al sector educativo que se están publicando en el portal de datos abiertos de Colombia, compararla con la disponibilidad de datos abiertos similares en otras iniciativas internacionales relevantes, analizar el potencial de apertura a corto plazo de nuevos conjuntos de datos relacionados con la cobertura y calidad educativa y analizar cómo se están utilizando y qué tipo de aplicaciones se están desarrollando con los datos abiertos en distintos países.

LA METODOLOGÍA

Este estudio sectorial parte del conocimiento sobre la preparación para la apertura de datos del conjunto del país, adquirido en el proceso de elaboración del informe ODRA¹.

La metodología sigue los siguientes pasos:



1 - Open Data Readiness Assessment: Evaluación de la Preparación sobre Datos Abiertos

El sector educación incluye a una amplia variedad de actores, que juegan papeles tanto en la oferta como en la demanda educativa. Cada actor tiene intereses diferentes y actúa en ejercicio de sus competencias, lo que constituye una red de relaciones compleja y rica. Las acciones que se planifiquen deben tener siempre en cuenta cuáles son los destinatarios, qué

valor aportan a cada uno y qué colaboración pueden aportar.

En las entrevistas mantenidas, se ha comprobado que existe un alto grado de alineamiento en cuanto a cuáles son los principales retos a abordar en el contexto actual. Algunos de estos retos tienen visos de poder beneficiarse de la apertura y uso de datos públicos.

PRINCIPALES RETOS EN EL SECTOR EDUCATIVO EN COLOMBIA CON ALTO DE POTENCIAL DE BENEFICIARSE DE LA APERTURA Y USO DE DATOS PÚBLICOS:

1 Entender la deserción y lanzar medidas para corregirla

4 Ligar calidad educativa a programas educativos

2 Ligar competencias a la empleabilidad

5 Mejorar las competencias docentes

3 Evaluar impacto de políticas de educación

6 Garantizar el buen uso de los recursos públicos en las instituciones educativas

EN CONCRETO, LOS DATOS ABIERTOS TIENEN EL POTENCIAL DE BENEFICIAR EL SECTOR EDUCATIVO EN COLOMBIA EN LOS SIGUIENTES ASPECTOS SEGÚN LA CATEGORÍA DE RETO:



1. RETO EN COBERTURA:

- A.** MONITORIZAR EL CUMPLIMIENTO DE UNA COBERTURA TOTAL
- B.** ESTRATEGIAS DE APOYO PARA LA PERMANENCIA
- C.** ESTRATEGIAS DE APOYO PARA MINIMIZAR ABSENTISMO
- D.** ENTENDER EL DESFASE ENTRE DEMANDA Y OFERTA
- E.** CARACTERIZACIÓN SOCIOECONÓMICA DE POBLACIÓN ESTUDIANTIL
- F.** SEGUIMIENTO DE MEJORAS EN INFRAESTRUCTURAS RURALES
- G.** DAR A CONOCER LA PLANTA DE PERSONAL
- H.** DAR A CONOCER LA EVOLUCIÓN DE LA MATRÍCULA
- I.** DAR A CONOCER LA EVOLUCIÓN DE ALUMNOS POR DOCENTE
- J.** DAR A CONOCER METROS CUADRADOS DISPONIBLES POR N° DE ALUMNOS
- K.** DAR A CONOCER LAS JORNADAS DISPONIBLES
- L.** ORIENTAR A LA INSTALACIÓN DE MAYOR CAPACIDAD DONDE SE PRECISE



2. RETO EN CALIDAD EDUCATIVA:

- A.** MEJOR COMPRENSIÓN DE LAS CAUSAS Y EFECTOS DE LA PROBLEMÁTICA EDUCATIVA
- B.** RESPUESTA RÁPIDA A MALOS RESULTADOS DETECTADOS
- C.** DAR SOPORTE A UN SISTEMA DE CALIDAD MULTIDIMENSIONAL
- D.** SOPORTE A LA GESTIÓN POR RESULTADOS
- E.** INFORMACIÓN PARA TOMA DE DECISIONES
- F.** ANÁLISIS SISTEMÁTICO DE LA POLÍTICA EDUCATIVA
- G.** INTEGRAR QUEJAS DE COMUNIDAD ESCOLAR EN LA GESTIÓN
- H.** RENDICIÓN DE CUENTAS Y CONTROL SOCIAL
- I.** DEMOSTRAR CON DATOS QUE LAS POLÍTICAS HAN SIDO (O NO) ÚTILES
- J.** INCENTIVO PARA MEJORAR LA INFORMACIÓN
- K.** ANÁLISIS DE PROGRAMAS EDUCATIVOS
- L.** PERSONALIZACIÓN DE PROGRAMAS EDUCATIVOS
- M.** SEGUIMIENTO DEL PROGRAMA DE BILINGÜISMO
- N.** ANÁLISIS DE INTERACCIÓN EN ENSEÑANZA MEDIA POR COMPUTADOR
- O.** ESTABLECIMIENTO DE INCENTIVOS A PARTIR DE MEJORES PRÁCTICAS
- P.** MEJORA DE COMPETENCIAS DOCENTES PARA MEJORES RESULTADOS



3. RETO EN GESTIÓN:

- A. MEJORA DE SERVICIOS PRESTADOS ONLINE A TRAVÉS DE LA ANALÍTICA WEB
- B. TRANSPARENCIA EN LA CONTRATACIÓN
- C. TRANSPARENCIA EN LA GESTIÓN ECONÓMICA
- D. TRANSPARENCIA EN LA GESTIÓN DE RRHH
- E. MONITORIZAR NIVEL DE EJECUCIÓN PRESUPUESTAL
- F. CONOCER LA INVERSIÓN POR ALUMNO
- G. MEJORAR CONSISTENCIA DE INFORMACIÓN REPORTADA
- H. CONOCER LA OPORTUNIDAD DE CUBRIR VACANTES DE DOCENTE
- I. CONOCER EL COSTO POR SERVICIO
- J. MEJORA DE PROCESOS DE ATENCIÓN CIUDADANA (P.E., CONVALIDAR TÍTULOS)
- K. IMPLICACIÓN COMUNIDAD EDUCATIVA EN GESTIÓN DE RECURSOS
- L. IMPLICACIÓN COMUNIDAD EDUCATIVA EN INSPECCIÓN DE SERVICIOS
- M. OPTIMIZACIÓN DEL USO DE RECURSOS FINANCIEROS



4. RETO EN EQUIDAD Y DESARROLLO:

- A. MEJORA DEL TRANSPORTE ESCOLAR
- B. MEJORA DE LA NUTRICIÓN ESCOLAR
- C. MEJORA DE LA SALUD ESCOLAR
- D. FACILITAR PROCESOS BUROCRÁTICOS A LOS USUARIOS
- E. SOPORTE A PROGRAMAS: EMBARAZO, DROGODEPENDENCIA, MALNUTRICIÓN...
- F. SEGUIMIENTO DE “SER PILO PAGA” Y OTROS PARA AUMENTAR IGUALDAD
- G. SEGUIMIENTO DE ENFOQUE DE GÉNERO
- H. PONER EN RELACIÓN COMPETENCIAS PROFESIONALES Y CURRÍCULO EDUCATIVO
- I. MEJOR ORIENTACIÓN AL EMPLEO
- J. ESTUDIOS DE PROSPECTIVA LABORAL



DATOS EN EL SECTOR EDUCATIVO EN COLOMBIA: OFERTA Y DEMANDA

EN cuanto a los datos ya publicados en el catálogo nacional de datos abiertos, existen 221 conjuntos de datos relacionados con el sector educativo, de 172 entidades distintas. Se ha detectado un alto nivel de duplicidad entre los datos publicados por las Alcaldías y los de MinEducación. En comparación con las mejores prácticas internacionales, se observa un desequilibrio en la tipología de datos publicados, ya que son mayoritariamente de infraestructuras educativas, con escasa presencia de otras categorías de datos.

En los distintos portales web de MinEducación se han identificado más de 600 archivos con datos, los cuales son potenciales *datasets* publicables a corto plazo como datos abiertos ya que éstos están actualmente disponibles para su descarga en formato XLS o CSV principalmente. La principal acción a corto plazo debe ser la publicación de estos datos en el portal de datos nacional.

MinEducación dispone de 115 sistemas de información, según la información facilitada por la Oficina de Tecnologías y Sistemas de Información. Estos sistemas cubren los ámbitos de educación básica, educación superior, planeación, Ministerio y otros de carácter interno. Se ha detectado que al menos 38 de ellos son susceptibles de ser fuentes de origen para abrir datos.

En cuanto a la demanda de datos por parte de la ciudadanía, ésta se limita a las instituciones de la sociedad civil que trabajan en este campo, principalmente con objetivos de investigación y activismo. El sector empresarial y los desarrolladores no conocen las necesidades del sector, por lo que apenas están tratando de aportar soluciones. Será importante conectar con ellos.

RECOMENDACIONES:

Entre las propuestas concretas, destacan las siguientes:

- Revisar las mejores prácticas internacionales aportadas en este informe, para inspirar nuevos usos de los datos.
- Dar los pasos para lograr un currículo educativo a lo largo de toda la vida del individuo, soportado en un identificador único que permita la anonimización de los registros
- Revisar las condiciones de uso de los portales de MinEducación, con avisos legales que permitan la reutilización para cualquier uso.
- Delimitar la información publicable por las Alcaldías y Gobernaciones para evitar duplicidades con la publicación de datos del Ministerio y entidades del orden Nacional.
- Detectar y difundir mejores prácticas locales en la creación de un ecosistema para los datos abiertos, como es el caso de la Secretaría de educación de Bogotá, incluyendo buenas prácticas en zona rural.
- Desarrollar un plan de comunicación de datos abiertos para educación con mensajes adaptados a los diferentes actores/reutilizadores de datos.
- Integrar en la planificación del sector algunas de las metas que plantea la sociedad civil, para aumentar la colaboración.
- Disponer datos de los contratos de servicios que realiza el Ministerio de Educación para el suministro de raciones alimenticias y poner las condiciones para que sirvan de base para mejorar la nutrición escolar.
- Determinar los orígenes de datos que tienen el mayor grado de actualización y precisión y utilizar éstas como fuente de datos principal.
- Detectar de manera temprana las oportunidades de uso de los datos del futuro marco nacional de cualificaciones.
- Involucrar a la Universidad para impulsar una experiencia de publicación de datasets en formatos semánticos (RDF).



A

METODOLOGÍA PARA APOYAR INICIATIVAS SECTORIALES DE DATOS ABIERTOS

A.1. ¿POR QUÉ REALIZAR UN ANÁLISIS SECTORIAL?

Las iniciativas de datos abiertos afrontan una irritable paradoja: pese a que existe un consenso respecto de su importancia, resulta muy difícil conseguir resultados relevantes. No podemos olvidar que la inversión dedicada a la apertura de datos no genera un retorno hasta que esos datos son utilizados, ya sea de manera interna o externa, ya sea para producir valor social o económico.

El caso es que resulta muy difícil visualizar posibles usos de los datos mientras mantenemos una visión general. En la realidad, los datos son útiles en el contexto de un sector concreto, habitado por actores reconocidos, que persiguen finalidades específicas.

Por lo tanto, la aproximación sectorial permite verticalizar una política tan transversal como es la de datos abiertos, para así extraer el mayor valor de ella. Esta visión vertical convive con la riqueza que ofrece la transversalidad a la hora de combinar datos de diferentes procedencias para el logro de unos fines determinados.

Una iniciativa de datos abiertos debe contemplar tanto la parte de la oferta de datos como de la demanda de su reutilización. Un enfoque sectorial, ofrece la oportunidad de enfocarse en los beneficios para segmentos de destinatarios bien identificados, como punto de partida para estimular el crecimiento de un ecosistema de reutilización de datos que ponga en juego la colaboración de agentes variados en la consecución de mayor transparencia, mejor toma de decisiones, nuevos servicios y el desarrollo de innovaciones que mejoren el bienestar ciudadano.

A.2. DIAGNÓSTICO ESTRATÉGICO COMO PUNTO DE PARTIDA

Este análisis sectorial comienza con el desarrollo de un diagnóstico del grado de preparación del país respecto de los datos abiertos, mediante el uso de la metodología ODRA. La herramienta ODRA utiliza un enfoque sistémico para abrir los datos, lo que significa que está diseñada para observar el entorno más amplio de datos abiertos. En el lado de la “oferta”, incluye asuntos como el marco de políticas y normas, los datos ya existentes en el gobierno, la infraestructura tecnológica y los estándares; mientras que el lado de la “demanda” surgen temas como los mecanismos de participación ciudadana, la demanda existente de datos del gobierno, y la existencia de comunidades de posibles usuarios.

El conocimiento del grado de preparación del país –o de la región, si fuera este el caso- proporciona un marco de comprensión necesario para poder realizar un análisis rápido de cada sector elegido.

Un primer paso será el de la elección del sector a estudiar. Se recomienda comenzar por un sector que ofrezca buenas perspectivas para alcanzar éxitos tempranos, ya sea por la alta disponibilidad de datos existente, o por la presencia de comunidades de usuarios con capacidad para la reutilización, o por ambas razones.

A.3. ANÁLISIS DEL SECTOR

EL METODOLOGÍA DE ANÁLISIS SECTORIAL SE ESTRUCTURA EN LOS SIGUIENTES NUEVE PASOS:

1. SELECCIÓN DEL SECTOR

Tal como se ha señalado, el paso previo para analizar un sector consiste en un primer análisis de oportunidades, para seleccionarlo de entre una lista corta. En este análisis se valora la disponibilidad de datos y otros aspectos estratégicos, a partir de la información manejada en el estudio ODRA. Este es un primer insumo que ayuda a comenzar el estudio.

En el caso de Colombia, el objetivo es identificar un sector sobre el cuál desarrollar una estrategia piloto para la dinamización de la apertura de datos en la administración del Gobierno de Colombia. La recomendación para la decisión del sector se centra en el objetivo de trabajar conjuntamente en un sector de alto impacto para la política pública de datos abiertos de Colombia con potencial suficiente para lograr un caso de éxito. Además, se busca que el sector seleccionado contribuya a evidenciar el valor económico y social a partir del aprovechamiento de los datos abiertos por parte de los usuarios y, de esta manera, sirva de ejemplo para el resto de entidades de gobierno. Los sectores candidatos son aquellos identificados prioritarios en el Mapa de Ruta del Plan de Acción 2015 de Gobierno en Línea.

De acuerdo a la información obtenida en las entrevistas durante la fase de trabajo de campo del ODRA y a las evidencias rescatadas de análisis de situación como el índice GEL para 2014, se determina que el sector mejor posicionado para el desarrollo de una acción específica de datos abiertos es Educación y donde concretamente, el Instituto Colombiano Para la Evaluación de la Educación (ICFES) destaca como un actor aliado muy relevante.

El sector de Educación, al tratarse de una política nacional, posee datos de alto interés público, sistemas de información robustos, una alta interacción con las entidades territoriales del sector y un grado inicial de publicación y apertura de datos destacable.

2. BÚSQUEDA DE INSPIRACIÓN

En segundo lugar, se analiza información secundaria acerca de las mejores prácticas sectoriales en el mundo, tanto respecto de la oferta de datos como de buenas prácticas reconocidas. Ese listado de buenas prácticas servirá de inspiración para imaginar futuros usos de los datos.

La recopilación del Catálogo de Prácticas Internacionales en Datos Abiertos del Sector de Educación como base para la inspiración se presenta al final de este documento (pág. 68).

3. ENTENDIMIENTO DE LAS METAS INSTITUCIONALES

Es vital que las propuestas que emanan del estudio tengan como referencia las metas establecidas por las instituciones públicas del sector. El resultado del análisis debe ser la elección de las temáticas clave en las que se va a enfocar en estudio. El hecho de elegir unos puntos focales no implica que se desestimen el resto de temas que puedan aparecer durante el estudio, sino que sirve para contribuir a una mayor economía de esfuerzos.

El análisis de las metas institucionales se presenta en la sección B (pág. 17) “La Educación en Colombia: Situación de Partida y Metas a Alcanzar”.

4. TRABAJO DE CAMPO

A continuación, se realizan entrevistas con agentes clave, tanto de la parte del Gobierno como de la sociedad civil y las empresas. Las entrevistas exploran la disponibilidad de datos y los beneficios que se perciben en su reutilización, junto con aspectos de contexto como las barreras y oportunidades que van a aparecer en la implantación de la iniciativa.

En el Anexo 1 se incluye la lista de instituciones y personas entrevistadas en Colombia durante la fase de trabajo de campo.

5. BARRERAS Y OPORTUNIDADES EN EL SECTOR

A partir de la información obtenida en las entrevistas, se recoge una lista de barreras y de oportunidades, para la apertura en el sector identificado que ofrecen indicios que ayudarán al planteamiento de recomendaciones para el plan de acción.

El análisis de barreras y oportunidades para la apertura de datos en el sector educativo en Colombia se presenta en la sección C (pág. 25) de este documento.



Estas barreras y oportunidades se clasifican en 6 áreas:

- | | |
|------------------------|----------------------|
| 1. LIDERAZGO | 4. TECNOLOGÍA |
| 2. LEGAL | 5. DATOS |
| 3. ORGANIZACIÓN | 6. DEMANDA |

A.4. INFORME FINAL

Los hallazgos y recomendaciones se reúnen en este informe final.

Este informe se ha sometido a discusión y validación, y las aportaciones recibidas se incluyen en este informe final.

6. BENEFICIOS DE LA REUTILIZACIÓN

Con la información disponible y a la luz de las mejores prácticas, se propone una lista de posibles beneficios a alcanzar mediante la reutilización de datos. En la medida de lo posible, estos posibles beneficios se ponen en relación con las metas sectoriales u otras referencias que agrupen los beneficios y mejoren su inteligibilidad.

Los beneficios de una Política de Apertura y Reutilización de datos se presenta en la sección D de este documento (pág. 37).

7 ANÁLISIS DE FUENTES DE DATOS PUBLICABLES

A partir de la información obtenida en las entrevistas y de la documentación recogida sobre los sistemas de información, se practica un primer levantamiento de posibles conjuntos de datos publicables. Este es un producto en sí mismo, que debe servir para que las autoridades tomen una decisión respecto de qué datos publicar y con qué prioridad.

El análisis de las fuentes de datos publicables del sector educativo en Colombia se presenta en la sección E de este documento (pág. 45).

8. ESTUDIO DE DEMANDA

El estudio de la demanda es clave en el análisis sectorial ya que permite determinar las necesidades de datos de la demanda de datos del sector, así como entre entidades constituyendo un espacio fundamental para la toma de decisiones y priorización de apertura de datasets.

El Estudio de Demanda del Sector Educativo en Colombia se presenta en la sección F de este documento (pág. 55).

9. RECOMENDACIONES PARA LA ACCIÓN

Como fruto de los análisis realizados, el informe concluye con una lista de recomendaciones para orientar un plan de acción sectorial. Las recomendaciones deben incluir acciones no sólo respecto de la publicación de datos, sino también elementos de las otras dimensiones que se contemplan en este estudio.

Las recomendación encaminadas a guiar un plan de acción se presentan en la sección G de este documento (pág. 62).





B

LA EDUCACIÓN EN COLOMBIA: SITUACIÓN DE PARTIDA Y METAS A ALCANZAR

B.1. SITUACIÓN DE COLOMBIA RESPECTO DE LOS PAÍSES DE LA OCDE

Aunque el alcance de este estudio no incluye una evaluación de las políticas educativas de la República de Colombia, el análisis incorpora alguna información que aporta contexto al reto que abordan las políticas educativas actuales, con la finalidad de entender el propósito de los esfuerzos que se realizan en el sector educación y, entre ellos, los que se pretenden iniciar por la línea de la apertura y reutilización de datos.

Para ello, se ha usado como referencia el informe de Educación En Colombia³ publicado por la OCDE en noviembre de 2016, y a partir de la “Revisión de las políticas nacionales de educación. La educación en Colombia”. Este estudio se construye en base a indicadores de cursos anteriores -básicamente, 2013 y 2014-, algunos de los cuales se están mejorando notablemente. Aun así, sirve de referencia para entender qué se está haciendo y por qué.

Estos son los grandes titulares del estado de la educación en Colombia, en comparación a los países de la OCDE, según la evaluación de 2016:

- A pesar del bajo nivel educativo de toda la población, una proporción importante de adultos tiene educación superior.
- A pesar de que el tiempo dedicado a la educación básica, media y terciaria se ha incrementado en dos años en solo una década, Colombia ofrece un promedio de la duración prevista de la educación más corta que la media de los países de la OCDE. Las diferencias entre hombres y mujeres en los campos de estudio entre los egresados universitarios son menos marcadas en Colombia que en la mayoría de los países de la OCDE.
- Colombia gasta en educación un porcentaje de su PIB superior al promedio⁴.
- La participación del financiamiento privado es de las más grandes entre los países de la OCDE.
- El gasto público por cada estudiante en la escuela es la mitad del correspondiente a cada universitario y equivale a un cuarto de lo invertido por el país promedio de la OCDE..
- Se matricula a menos niños que el promedio OCDE en la educación inicial, especialmente en niños de 3 años.
- Aquellos estudiantes que siguen en el sistema con 15 años tienen un desempeño unos tres años por debajo del de los estudiantes de la misma edad de los países de la OCDE en PISA.
- Una proporción de estudiantes superior al promedio se inscribe en instituciones privadas.
- Tomando como referencia los jóvenes de 15 años, las escuelas privadas tienden a propiciar más el aprendizaje, al aplicar un mayor volumen de recursos.
- El rendimiento de los estudiantes se relaciona fuertemente con las características socioeconómicas de los alumnos y de las escuelas.
- En educación terciaria, solo el 9% de las familias más pobres logran matricularse comparado con el 53% de las familias más ricas.
- Los estudiantes de primaria y de secundaria asisten a la escuela más días por año que el promedio de la OCDE.

3 - <https://www.oecd.org/education/school/Educacion-en-Colombia-Aspectos-Destacados.pdf>

4 - En 2011, Colombia gastó el 6,7% de su PIB en educación (fuentes públicas + privadas), mientras que el promedio de la OCDE es el 6,1%.

Estos grandes titulares se relacionan con la siguiente tabla de indicadores clave.

REVISIÓN DE POLÍTICAS NACIONALES DE EDUCACIÓN. LA EDUCACIÓN EN COLOMBIA 2016 (OCDE)

INDICADORES CLAVE	COLOMBIA	PROMEDIO O TOTAL	MIN. OCDE	MAX. OCDE
INFORMACIÓN SOBRE LA EDUCACIÓN				
Duración de la educación obligatoria	10	10	8	12
Edad a la que empiezan a estudiar (OCDE, 2014a)	5	6	4	7
Esperanza de vida escolar, número estimado de años desde primaria hasta educación superior, 2011 (UNESCO -UIS, 2015)	13,5	16,5	12,8	19,9
Primera edad de selección en el sistema educativo (OCDE, 2013b)	15	14	10	16
TASAS DE MATRICULADOS				
Tasas de matriculados entre 3 y 4 años de edad en educación para la primera infancia expresados como porcentaje de la población del mismo grupo etario, 2012 (OCDE, 2014a)	61%	76%	12%	99%
Rango de edades en el cual más del 90% de la población en edad escolar está matriculada, 2012 (OCDE, 2014a)	7-13	4-16	6-13	3-18
Índices brutos de matriculados en educación superior, 2013 (UNESCO-UIS, 2015)	48%	71,5%	29%	116%
% de estudiantes que reportan que han repetido por lo menos un año en primaria, básica secundaria o educación media (OCDE, 2013b)	40,6%	12,4%	0,0%	36,1%
FINANCIACIÓN				
% de estudiantes que reportan que han repetido por lo menos un año en primaria, básica secundaria o educación media (OCDE, 2013b)	40,6%	12,4%	0,0%	36,1%
Gasto anual por estudiante por institución educativa, para todos los servicios, en USD equivalentes convertidos usando la PPA del PIB, 2011 (OCDE, 2014a)				
Educación preprimaria	3.491	7.428	2.412	25.074
Educación primaria	2.041	8.296	2.218	23.871
Educación secundaria	2.207	9.280	2.736	16.182
Educación superior	6.882	13.958	7.868	26.021
Proporciones relativas de gasto público y privado en instituciones educativas, 2011 (OCDE, 2014a)				
Recursos públicos	65.1%	83,9%	59,9%	97,6%
Todas las fuentes privadas	34,9%	16,1%	2,4%	40,1%

INDICADORES CLAVE	COLOMBIA	PROMEDIO O TOTAL	MIN. OCDE	MAX. OCDE
-------------------	----------	------------------	-----------	-----------

Resultados educativos

Tasas de niveles de estudios alcanzados, 2013 (OCDE, 2015b)

% de personas entre 25 y 34 años de edad cuyo nivel de estudios alcanzados es educación básica secundaria o inferior	34,0%	17,0%	2,0%	52,0%
% de personas entre 25 y 34 años de edad cuyo nivel de estudios alcanzados es por lo menos educación media	40,0%	43,0%	23,0%	65,0%
% de personas entre 25 y 34 años de edad cuyo nivel de estudios alcanzados es educación superior	26,0%	41,0%	22,0%	67,0%
% de personas entre 15 y 29 años de edad que ni estudian, ni trabajan, ni reciben formación, 2013 (OCDE, 2015b)	21,6%	15,5%	6,1%	31,3%

Tasas de desempleo de las personas entre 25 y 34 años de edad por nivel de estudios alcanzados, 2013 (OCDE, 2015b)

Por debajo de educación media	8,6%	20,6%	5,0%	54,0%
Educación media y educación postsecundaria no universitaria	11,0%	10,9%	3,5%	37,1%
Educación superior	10,5%	7,7%	2,8%	33,1%

NÚMERO DE SEDES ESCOLARES Y ESTUDIANTES (2014)

	SEDES ESCOLARES		ESTUDIANTES		
	Educación Pública	Educación Privada	Educación Pública	Educación Privada	Total
TOTAL	44.416	6.575	6.434.700	1.065.987	7.500.687
POR NIVEL DE EDUCACIÓN					
Básica primaria	43.739	6.399	3.735.902	640.822	4.376.724
Básica secundaria	10.814	3.418	2.698.798	425.165	3.123.963
POR UBICACIÓN					
Urbana	9.718	6.256	4.568.827	1.007.458	5.576.285
Rural	34.698	319	1.865.873	58.529	1.924.402

PERSONAL ACADÉMICO, POR NIVEL EDUCATIVO (2010 Y 2013)

Nivel Educativo	2010		2013	
	CANTIDAD	PORCENTAJE	CANTIDAD	PORCENTAJE
Pregrado	34.026	32,4%	43.801	37,5%
Especialización	37.496	35,7%	39.556	33,9%
Maestría	27.048	25,7%	26.668	22,8%
Doctorado	6.595	6,3%	6.794	5,8%
Total	105.165	--	116.819	--

Nota: según datos preliminares de MEN-SNIES del 2013; hasta el 10 de mayo del 2013

B.2. ANÁLISIS DEL SECTOR

1. SELECCIÓN DEL SECTOR

Cada uno de los titulares, con sus aspectos positivos y negativos, está apuntando a cuatro grupos de metas a alcanzar en educación:



COBERTURA

Conseguir un adecuado encaje entre la oferta y la demanda educativa, mediante la adaptación de la capacidad instalada a las necesidades locales.



CALIDAD

Implantación de sistemas de medición de indicadores pedagógicos, junto con las acciones para la mejora continua.



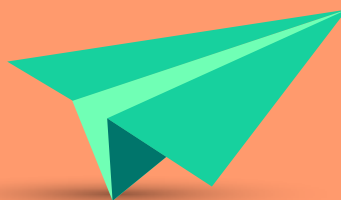
GESTIÓN

Aumentar la eficiencia del plan de desarrollo educativo y de la satisfacción de la comunidad educativa.



EQUIDAD

Atender a la población vulnerable para que alcance la igualdad de oportunidades durante el período educativo.



B.3. METAS DE COLOMBIA PARA EL SECTOR EDUCACIÓN

En el Plan Nacional de Desarrollo (PND) 2015-2018 se plantean una serie de metas a alcanzar, con su temporalización. Es necesario entender estas metas, para entender la contribución, directa e indirecta, que la apertura y reutilización de datos puede aportar al cumplimiento de estas metas. 2014 (en un caso, 2013) marca la línea base para avanzar.

Existen unas metas trazadoras, a manera de índices fundamentales:

METAS TRAZADORAS	#	INDICADOR	2014	CUATRIENIO
	1	Porcentaje de colegios oficiales en las categorías A+, A y B en las pruebas Saber 11	33,3%	45%
	2	Tasa de cobertura de alta calidad en educación superior (Instituciones de Educación Superior o programas acreditados)	14,6%	20%
	3	Porcentaje de estudiantes evaluados en el sector oficial en las pruebas SABER 5 que sube de nivel de logro, respecto a la línea base	0 %	12%
	4	Porcentaje de estudiantes del sector oficial evaluados con nivel B1 o superior de inglés del Marco Común Europeo	2,26%	8%
	5	Porcentaje de estudiantes con jornada única	0%	30%
	6	Tasa de cobertura bruta en educación media	77,3%	83%

El PND marca también una lista de metas intermedias y productos:

AMPLIACIÓN DE COBERTURA CON CIERRE DE BRECHAS	#	META	2014	CUATRIENIO
	7	Porcentaje de estudiantes evaluados en el sector oficial en las pruebas SABER 5 que sube de nivel de logro, respecto a la línea base	0 %	12%
	#	PRODUCTO	2014	CUATRIENIO
	8	Porcentaje de estudiantes en establecimientos focalizados por el Programa Todos a Aprender en niveles Satisfactorio y Avanzado en pruebas de lenguaje SABER 5	27,7%	30,9%
	9	Porcentaje de estudiantes en establecimientos focalizados por el Programa Todos a Aprender en niveles Satisfactorio y Avanzado en pruebas de matemáticas SABER 5	21,7%	29,9%
	10	Número de docentes por tutor en el programa Todos a Aprender	27	20
	#	META INTERMEDIA	2013	CUATRIENIO
	11	Tasa de supervivencia de grado primero a once	37,2%	45,0%
	#	PRODUCTO	2013	CUATRIENIO
	12	Porcentaje de matrícula oficial con conexión a internet	68,0%	90%
	13	Proporción de niños entre 6 y 16 años en el hogar que asisten al colegio	87,1%	93%
	14	Tasa de deserción intra-anual de educación preescolar básica y media	3,1%	2,5%
	#	META INTERMEDIA	2014	CUATRIENIO
	15	Tasa de cobertura bruta en educación media en la zona rural	62,1%	69,0%
	#	PRODUCTO	2014	CUATRIENIO
	16	Sedes rurales intervenidas con mejoramiento o construcción de infraestructura	168	1.000

EXCELENCIA DOCENTE	#	META INTERMEDIA	2014	CUATRIENIO
	17	Porcentaje de docentes que entran al magisterio que se encuentran en los quintiles superiores (4 y 5) de las pruebas SABER PRO en Razonamiento Cuantitativo	37,1%	45%
	18	Porcentaje de docentes que entran al magisterio que se encuentran en los quintiles superiores (4 y 5) de las pruebas SABER PRO en Lectura Crítica	35,7%	42%
	#	PRODUCTO	2014	CUATRIENIO
	19	Estudiantes beneficiados con créditos condonables para programas profesionales de licenciatura en Instituciones Educativas Certificadas con alta calidad	117	2.000
	20	Porcentaje de programas de licenciatura con acreditación de alta calidad	16,5%	90%
	21	Porcentaje de estudiantes de licenciatura en nivel de desempeño alto en pruebas de razonamiento cuantitativo de SABER PRO	3,4%	13,2%
	22	Porcentaje estudiantes de licenciatura en nivel de desempeño alto en la prueba SABER PRO en lectura crítica	9,3%	14,7%
	23	Porcentaje de personas que ingresan a programas de licenciatura que están entre los puestos del 1 a 400 de la prueba SABER 11	53,1%	75%
	#	META INTERMEDIA	2014	CUATRIENIO
24	Porcentaje de Docentes del sector oficial de educación preescolar, básica y media con formación de Postgrado	31,6%	38%	
#	PRODUCTO	2014	CUATRIENIO	
25	Número de docentes beneficiados con recursos de la nación para mejorar su nivel de formación a nivel de postgrado	-	17.000	

EDUCACIÓN SUPERIOR	#	META INTERMEDIA	2014	CUATRIENIO
	26	Porcentaje de cupos de educación técnica y tecnológica con acreditación de alta calidad en programas e IES con acreditación de alta calidad	10,2%	15%
	#	PRODUCTO	2014	CUATRIENIO
	27	Nuevos cupos en educación técnica y tecnológica	147.729	150.000
	28	Tasa de deserción anual en educación técnica y tecnológica	19,43%	15,0%
	#	META INTERMEDIA	2014	CUATRIENIO
	29	Tasa de Cobertura en Educación Superior	46,1%	57%
	#	PRODUCTO	2014	CUATRIENIO
	30	Nuevos Cupos en Educación Superior	464.164	400.000
	31	Departamentos con tasa de Cobertura en Educación Superior por encima del 20%	25	33
32	Tasa de deserción en educación superior	10,1%	8,0%	
#	META INTERMEDIA	2014	CUATRIENIO	
34	Estudiantes beneficiados con nuevos créditos condonables	26.899	125.000	
35	Porcentaje de créditos nuevos de pregrado aprobados en Programas o Instituciones Educativas Certificadas con alta calidad	39%	60%	
36	Docentes de Educación Superior con formación Doctoral	7.573	10.000	
37	Ganancia en puestos de las Instituciones de Educación Superior colombianas en los ranking internacionales	0	25	

COLOMBIA LIBRE DE ANALFABETISMO	#	META INTERMEDIA	2014	CUATRIENIO
	38	Tasa de analfabetismo para población de 15 años y más		5,81%
	#	PRODUCTO	2014	CUATRIENIO
	39	Nuevos jóvenes y adultos alfabetizados	419.082	676.000

COLOMBIA BILINGÜE	#	META INTERMEDIA	2014	CUATRIENIO
	40	Docentes de inglés del sector oficial evaluados con nivel B2 o superior de acuerdo a los niveles del Marco Común Europeo de Referencia.		0%
	#	PRODUCTO	2014	CUATRIENIO
	41	Docentes formados en inglés	5.703	8.000





C

BARRERAS Y OPORTUNIDADES PARA LA APERTURA DE DATOS EN EL SECTOR EDUCATIVO

LA Educación en Colombia es un sector prioritario para el país, que está experimentando un rápido desarrollo, en base a las metas expuestas en el Plan Nacional de Desarrollo. En todo caso, este estudio no pretende evaluar el estado actual de la Educación en Colombia, ni mucho menos realizar un análisis comparativo con otros países.

Por lo tanto, la lista de barreras y oportunidades que figuran en este capítulo debe ser entendida como una búsqueda de elementos de análisis que nos permitan lanzar recomendaciones que ayuden a una mejor implantación de la política de datos abiertos en el sector. Se ha puesto un mayor foco en las barreras que en la descripción de puntos fuertes y oportunidades, puesto que aquellas ofrecen una información más útil para los propósitos de este estudio.

Definimos así barreras y oportunidades:

BARRERAS: elementos que conviene mejorar o superar para ayudar a la iniciativa de datos abiertos

OPORTUNIDADES: espacios de posibilidad que surgen en torno a la dinámica educativa en el país.

C.1. ACTORES CLAVE DEL SECTOR

El sector educación incluye a una amplia variedad de actores, que juegan papeles tanto en la oferta como en la demanda educativa. Sin ser exhaustivos podemos distinguir:

- Planificadores/financiadores/Evaluadores del sistema: MinEducación, ICFES
- Otros entes públicos relacionados: DANE, MinSalud, MinEmpleo
- Entes locales: Secretarías de educación, Centros de Educación Superior
- Establecimientos educativos a distinto nivel
- Rectores de establecimientos
- Docentes de los distintos niveles educativos
- Investigadores del área
- Sociedad civil
- Sector empresarial
- Padres y alumnos
- Comunidad educativa
- Gobierno y entes multilaterales de desarrollo

En el caso de este estudio, hemos conseguido un amplio acceso a instituciones tanto de la oferta como de la demanda educativa. En el anexo 1 se ofrece la lista completa de entrevistas.

C.2. BARRERAS PARA LA POLÍTICA DE APERTURA Y REUTILIZACIÓN DE DATOS

1. DIFICULTAD DE GESTIONAR LA DIVERSIDAD DE ACTORES Y DE VISIONES

Aunque entendemos la diversidad de miradas sobre el campo de la educación como una riqueza neta y una oportunidad para lograr metas estratégicas complejas, también hay que considerar que esta diversidad de intereses y de lógicas, al aumentar la complejidad del contexto, constituye por sí misma un reto para los gestores públicos. No es fácil coordinar una red tan extensa y no es posible satisfacer plenamente a todas las partes. Se requiere una visión clara, compartida por la mayor parte del sector, para que los agentes puedan sumarse a la apertura de datos de la manera más autónoma posible. Se requiere, además, un buen sistema de información que mantenga a cada actor local informado de lo que sucede de manera global, para aprovechar las buenas prácticas en el conjunto de la red.

La educación en Colombia se caracteriza por permitir grados de libertad en el programa educativo de cada centro, establecido en el propio Plan Estratégico Institucional (PEI). Cada colegio tiene un PEI que hace que se oriente hacia una materia determinada. Este enfoque, muy interesante desde el punto de vista didáctico y de producción de innovación educativa, añade complejidad al sistema de información.

2. OPEN DATA ES UNA POLÍTICA NO SUFICIENTEMENTE IMPLANTADA

Aunque MinEducación y el resto de entidades públicas del sector están publicando datos en el portal nacional, la inmensa mayoría de información aún no se publica de esta manera. En muchos casos, se trata de información ya descargable como datos desde los portales Educación y que podría incorporarse con suma facilidad al portal nacional datos.gov.co.

Más allá del sector educativo, este es un caso particular del problema ya detectado en el informe ODR global de país. Se está necesitando reeditar el compromiso de cada institución con la apertura de datos en el portal nacional, a partir del levantamiento de datos en cada una de ellas.

Otro elemento a destacar es la inexistencia de acceso programático, vía API o similar, a los repositorios de datos de MinEducación, lo que limita las posibilidades de reutilización.





3. INCENTIVOS PERVERSOS DE LOS SISTEMAS DE MEDICIÓN

Toda medición introduce un marco de incentivos que favorece que el agente evaluado adecue su comportamiento al estándar premiado. Educación está implantando distintos sistemas de evaluación que, junto con sus obvios efectos positivos, están provocando efectos no deseados que es preciso controlar. Esto conduce a cierto grado de escepticismo respecto de los índices oficiales.

Así, la política que pone en relación al presupuesto de los establecimientos educativos con el número de alumnos, ha motivado, por ejemplo, el fenómeno de los “niños fantasma”; esto es, se reportan estudiantes inexistente para obtener mayor presupuesto de gestión del centro.

El índice sintético de calidad de los centros, que supone un excelente trabajo para la emulación y la mejora de cada centro, tampoco escapa de este efecto. Una forma de puntuar mejor es animar a que los peores alumnos abandonen el centro, lo que introduce problemas de inequidad. Por otra parte, puede producir desánimo, si no va acompañado de un acompañamiento cercano al centro, para ayudarles en la mejora.

También se reportan problemas políticos y sindicales fruto de la comparación entre centros. Estos efectos no deseados no deben servir de excusa para dejar de medir y de publicar los resultados, sino que representan riesgos que se deben prever para mitigarlos.

4. ESTRICTA APLICACIÓN DE LA NORMATIVA DE PROTECCIÓN DE DATOS

Colombia ha optado por un modelo de alta protección de los datos personales. La normativa existente limita de manera exigente los posibles usos de los datos que contengan alguna información de carácter personal. Ahora bien, hemos encontrado evidencias de que no sólo la ley es restrictiva, sino que además se tiende a aplicar con aún más intensidad de la estrictamente necesaria.

Cada vez que se propone liberar un dato, lo primero que se plantea es la aplicación de un filtro jurídico, en un contexto de no estandarización previa de los casos que se permiten o prohíben. Esto es, se tiende a colocar un análisis de habeas data caso a caso, en vez de realizar una distinción previa de qué tipologías de datos son publicables son restricción, de manera que sólo haya que filtrar otros datos más sensibles.

5. FALTAN SISTEMAS QUE RELACIONEN LA EDUCACIÓN Y EL EMPLEO

El sector educación no ha establecido suficientes lazos de relación con el sector empleo, pese a la evidente oportunidad que ofrece. Especialmente, se echa en falta una mirada conjunta a las competencias profesionales de la educación superior, desde un punto de vista de información. El Servicio Nacional de Empleo y MinEducación pueden encontrar un campo de colaboración fructífero en este campo.

El Servicio Nacional de Aprendizaje (SENA) es una institución relativamente desgajada del sector educativo. A muchos efectos, el alumno que ingresa en el sistema SENA abandona el sistema educativo para integrar el sistema de empleo. Esto está dificultando la posibilidad de ligar datos y resultados entre la educación y el empleo.

6. DESEQUILIBRIO EN DOTACIÓN TECNOLÓGICA Y COBERTURA DE COMUNICACIONES

Colombia es un país extenso y muy variado, con extremas dificultades orográficas para la extensión de las comunicaciones a todos los puntos del país. Paralelamente, cuenta con un elevado índice de desigualdad socioeconómica, ligada, entre otros factores, a lo geográfico. De ahí que veamos municipios, como el de Bogotá, donde se han implantado excelentes prácticas de apropiación y uso de la tecnología, frente a otras zonas donde es complicado alcanzar los estadios mínimos de sociedad de la información.

7. AUSENCIA DE UN EXPEDIENTE EDUCATIVO UNIFICADO

MinEducación recoge mucha información individual de los estudiantes en Colombia, por ejemplo las que alimentan las estadísticas de ausentismo y deserción, junto con las de calidad. Sería deseable poder trazar esa información a lo largo de toda la vida de los individuos, para poder entender mejor las causas de los problemas y los efectos de las políticas emprendidas. De hecho, es una información que se solicita por parte de la sociedad civil para la realización de sus investigaciones.

El currículo único enfrentaría problemas legales a la hora de abrir los datos, que en buena medida se resolverían mediante la anonimización de la información. Ahí encontramos la dificultad añadida de la carencia de un código de identificación individual a lo largo de toda la vida. Una persona se identifica con hasta 3 códigos distintos a lo largo de su transcurso por el sistema educativo, lo que impide el seguimiento de historias individuales. Este problema se está corrigiendo para las generaciones que comienzan ahora, pero aún toca trabajar con datos que no poseen este identificativo único.

Un ejemplo de la dificultad de pensar en expedientes individuales lo encontramos en ICFES, que publica sólo resultados agregados de sus pruebas de grado 3, 5, 7 y 9, mientras que publica resultados individuales de nivel 11.





8. BAJA CALIDAD DE LOS DATOS EN ORIGEN

Como consecuencia de la falta de cultura del reporte, y de otros factores, la calidad del dato en origen no siempre es la adecuada. Se han dado algunos episodios de crisis al publicarse reportes que no correspondían con la calidad percibida por los rectores, por ejemplo en el ámbito universitario y cuya causa estaba en la mala cumplimentación de la información solicitada por parte de los productores; esto es, en este caso, de las propias Centros de Educación Superior.

Se están emprendiendo medidas prometedoras para reducir este problema, como el establecimiento de sistemas automáticos que rechazan el registro de información si detectan inconsistencias en la información. Ahora bien, no sólo son este tipo de datos los que deben mejorar su recogida. Por ejemplo, también se reporta baja calidad de los datos del sistema de Recursos Humanos. En este y en otros casos, la información se actualiza de manera tardía.

Se ha detectado otro aspecto de la baja calidad del dato en la existencia de fuentes múltiples de un mismo ítem de información, no coincidentes entre sí. Esto se debe a la carencia de fuentes de información únicas bien controladas.

9. FALTA DE ESTÁNDARES EN LA INFORMACIÓN EDUCATIVA

Se ha encontrado baja producción y adopción de estándares y de otros elementos de formalización de la información. El Directorio Único Educativo (DUE) es uno de los casos en donde la formalización sería de gran ayuda para poder producir datos que unan geoposición e información sobre los servicios. En otros casos, la información ni siquiera se presenta de manera estructurada, sino a manera de informes discursivos, Buena parte de los sistemas de rendición de cuentas adolecen de este problema.

Se reporta, asimismo, escasez de recursos para las actividades de procesamiento previas a la publicación, lo que redundando en este problema y afecta a la baja calidad de información.

Se echan en falta esquemas de interoperabilidad. Incluso, no existe una clasificación de la información que se solicita en el servicio de atención ciudadana, lo que impide alcanzar una comprensión de qué información debería abrirse como datos de manera prioritaria.



10. DIFICULTAD PARA LA SOCIEDAD CIVIL DE ACCEDER A DATOS

Aunque MinEducación publica mucha información, la sociedad civil reporta de manera unánime la dificultad que encuentran para acceder a la información que realmente necesitan. No se comparten datos que dan soporte a las políticas, o no se actualizan con suficiente prontitud. Cada organización de la sociedad civil establece su propia red de relaciones personales para la solicitud de información, lo que resulta ineficiente y, generalmente, insatisfactorio. Muchos datos de consumo interno podrían abrirse, en las condiciones de seguridad y protección de la intimidad adecuadas.

Curiosamente, la ley de acceso a la información ha llevado a que se cierren algunos repositorios de datos que estaban abiertos y cuyos datos ahora hay que solicitar por el canal establecido para la petición de información.

11. CARENCIAS EN LA INFORMACIÓN SOBRE INFRAESTRUCTURAS

Aunque existe información acerca de establecimientos educativos, no se ha consolidado un inventario de infraestructuras educativas. Este sería un prerrequisito para poder instalar sistemas de información geográfica que unan la localización de los centros a la información de servicio y a la información de rendición de cuentas.

12. EL MERCADO NO CONOCE BIEN LA PROBLEMÁTICA DEL SECTOR EDUCATIVO

A la hora de buscar una colaboración público-privada en la búsqueda de soluciones o de nuevos servicios, faltan empresas de desarrollo que entiendan bien la problemática del sector. Seguramente, sea MinEducación quien debería dar el paso de compartir su necesidades con los sectores empresariales que pueden aportar soluciones.

La reutilización de datos debería ser una de las prácticas clave de esta colaboración público privada. A día de hoy, las apps de Educación recaen sobre el área TIC del Ministerio, pero no se ha trabajado para conseguir un ecosistema de reutilización. No existe impulso a la iniciativa privada en el desarrollo de aplicaciones.

13. CARENCIA DE MECANISMOS DE COLABORACIÓN PARA LA MEJORA DE DATOS

No existe una relación formal entre MinEducación y las entidades de la sociedad civil para la colaboración en la mejora de la calidad de los datos ni para la producción de nuevos datos. Aquí se está desaprovechando un potencial que podría ser de gran ayuda. Desde luego, la apertura de datos permitiría una colaboración masiva.

Resumen de barreras detectadas

#	BARRERA	ÁREA
1	Dificultad de gestionar la diversidad de actores y de visiones	1. Estrategia
2	Open data es una política no suficientemente implantada	1. Estrategia
3	Incentivos perversos de los sistemas de medición	1. Estrategia
4	Estricta aplicación de la normativa de protección de datos	2. Legal
5	Faltan sistemas que relacionen la educación y el empleo	3. Organización
6	Desequilibrio en dotación tecnológica y cobertura de comunicaciones	4. Tecnología
7	Ausencia de un expediente educativo unificado	5. Datos
8	Baja calidad de los datos en origen	5. Datos
9	Falta de estándares en la información educativa	5. Datos
10	Dificultad para la sociedad civil de acceder a datos	5. Datos
11	Carencias en la información sobre infraestructuras	5. Datos
12	El mercado no conoce bien la problemática del sector educativo	6. Demanda
13	Carencia de mecanismos de colaboración para la mejora de datos	6. Demanda



C.3. OPORTUNIDADES PARA LA POLÍTICA DE APERTURA Y REUTILIZACIÓN DE DATOS

1. DIVERSIDAD DE ACTORES Y COMPLEJIDAD DEL CAMPO EDUCACIÓN

La diversidad de actores y de necesidades en el campo de la educación supone una riqueza neta y una oportunidad para lograr metas estratégicas complejas. Esta oportunidad sólo se pone en valor si se consigue articular una red de relaciones, soportada en sistemas de información, para combinar información de diferentes áreas.

Educación ya ha establecido convenios para compartir datos con otras instituciones como salud, SENA, transportes o DANE. Hasta ahora, se han planteado de manera individual, con servicios web ad hoc para cada caso, lo que motiva cierta dispersión e incluso un desconocimiento de cuántos servicios web hay establecidos. La apertura de datos vendría a simplificar enormemente este panorama y abriría mayores posibilidades a la colaboración interna.

Educación participa en el piloto nacional de interoperabilidad, que pretende poner las bases para un intercambio de información fiable y eficiente.

2. EXISTENCIA DE UNA POLÍTICA NACIONAL DE APERTURA DE DATOS

Colombia ocupa un lugar destacado entre las naciones que abren datos en el mundo. Está afrontando en estos momentos un proyecto de relanzamiento de su iniciativa, a partir del diagnóstico ODRA realizado en colaboración con el Banco Mundial. En breve, va a afrontar la migración a una nueva plataforma de publicación de datos que facilitará la publicación de datos de cada sector, junto con herramientas para un mejor consumo y una mejor presentación.

Educación no puede dejar pasar la oportunidad de ser uno de los mayores contribuyentes al portal nacional de datos y recolectar así los beneficios de compartir información con otras áreas de Gobierno en un espacio abierto a la colaboración.

3. MANEJO DE INCENTIVOS DE LOS SISTEMAS DE MEDICIÓN

Las mediciones de calidad de los centros, de los docentes y de los rectores introducen sistemas de incentivos complejos, que se están gestionando mediante programas de mejora. A medida que avanza la cultura del reporte, se producen pequeñas crisis que, cuando son bien gestionadas, ayudan a saltar escalones en el proceso de cambio. Uno de estos casos fue el de los resultados del primer informe MIDE, donde algunas instituciones de educación superior aparecieron en puestos inferiores del ranking debido a la poca atención que dedicaron a la cumplimentación de los cuestionarios, tal como se mencionó al hablar de las barreras. A medida que los índices se hacen más populares, aumenta la presión para obtener buenas puntuaciones en ellos.

4. BUENAS PRÁCTICAS LOCALES

El nivel de desarrollo de las políticas educativas y de la adopción de la tecnología varía ampliamente dentro del país. En positivo, hay que señalar que algunas secretarías de Educación ofrecen excelentes prácticas de las que aprender.

La Secretaría de Educación de Bogotá es una de estas. Por ejemplo, cuentan con SISCAP, sistema que ofrece información posicionada, en tiempo real, de la agenda y recursos de cada centro. La dotación de recursos tecnológicos en este caso es también óptima y viene acompañada de programas educativos innovadores, para sacarles el mayor provecho.

5. SE HAN INICIADO TRABAJOS PARA UNA MEJOR RELACIÓN CON LA INFORMACIÓN DE EMPLEO

SENA ha lanzado un “Observatorio Laboral y Ocupacional” que proporciona información estadística del monitoreo de las ocupaciones del país, las normas de competencia por ocupación, los programas más sobresalientes del SENA y las principales variables económicas de una región.

Por otra parte, se cuenta con “Graduados Colombia”, observatorio laboral para la educación. Es un sistema de información que brinda herramientas valiosas para analizar la pertinencia de la educación a partir del seguimiento a los graduados y su empleabilidad en el mercado laboral. De esta manera, contribuye al mejoramiento de la calidad de los programas académicos ofrecidos.

El Ministerio de Empleo está desarrollando el Marco Nacional de Cualificaciones, que será clave para poder poner en relación, por ejemplo, la oferta y demanda de competencias en el ámbito laboral. La constitución de la Mesa Nacional de Capital Humano ofrece oportunidades para avanzar en este ámbito.



6. PROGRAMAS INTENSIVOS DE EVALUACIÓN DE LA CALIDAD EDUCATIVA

Las pruebas SABER y SUPÉRATE, extendidas en todo el sistema educativo público, con tomas de información cada dos años, y con programas de seguimiento cercano, suponen una enorme riqueza de información que en parte ya está abierta, pero que debe ser publicada con lógica de dato abierto.

Lo mismo puede decirse de los sistemas de evaluación de establecimientos educativos, de rectores y docentes. El futuro portal de calidad educativa del ICFES debería orientarse a resolver los retos detectados y a servir como base para la federación de datos con el portal nacional datos.gov.co.

7. RENOVACIÓN DE SISTEMAS DE INFORMACIÓN

Hemos señalado, como barrera, la obsolescencia de los sistemas de información de MinEducación. En el apartado de oportunidades debemos referirnos a los múltiples proyectos puestos en marcha para su modernización.

Destacan dos proyectos especialmente:

Proyecto RENE (Registro Nacional de Educación)
Proyecto BI (Business Intelligence)

El número de sistemas de información existente es alto –según la información facilitada, 115–, lo que constituye una riqueza de información muy destacable. ICFES dispone de una bodega de datos para las evaluaciones SABER. Otros sistemas similares están apareciendo.

8. EXISTENCIA DE APLICACIONES A PARTIR DE DATOS

El Ministerio de Educación posee experiencia en la implementación de aplicaciones a partir de datos abiertos, desarrolladas en colaboración con MINTIC:

- Buscando colegios, sistema de consulta de instituciones educativas del país. Con datos basados en autoevaluaciones.
- Buscando carrera
- Sistema de atención al ciudadano (SAC)



9. SISTEMAS PARA LA MEJORA DEL REPORTE DE DATOS

MinEducación está abordando una serie de medidas para mejorar la calidad de los reportes y el grado de actualidad de los datos. Los datos se reportan de manera electrónica e incluso algunos de los procesos de obtención del dato –las pruebas SABER– se realizan de manera electrónica, lo que garantiza la exactitud del dato.

Se están implantando sistemas automáticos de filtrado que rechazan los registros defectuosos, antes de enviarlos. No se permite enviar el paquete de datos mientras exista algún error detectado. Además se están realizando auditorías de datos, no sólo para corregir problemas de calidad, sino también para evitar información fraudulenta.

10. MUCHOS DATOS YA DISPONIBLES

Pese a los problemas de calidad detectados, hay que señalar que los datos publicados en el portal del Ministerio son útiles y, en general, fiables. La calidad disminuye a medida que nos alejamos de la información tutelada de manera más cercana por el Ministerio.

MinEducación e ICFES publican datos en el portal nacional, pero aún muchos más en sus respectivos portales, o también en servidores FTP, para usuarios autorizados. En este informe se realiza un análisis más detallado de los datos disponibles.

La mayoría pertenecen a uno de estos cuatro tipos:

- Establecimientos
- Programas
- Matrículas
- Estadísticas





11. RIQUEZA DE DATOS ESTADÍSTICOS

Hay que hacer un aparte para señalar la riqueza de información estadística publicada. No es fácil encontrarla y, en ocasiones, es preciso acceder con clave de usuario o previa petición. Sin embargo, la producción de información es muy alta.

Está en marcha un proceso de identificación de necesidades conjuntas con el DANE para profundizar en la producción estadística.

12. ACERCAMIENTO A LAS ORGANIZACIONES DE LA SOCIEDAD CIVIL

El Ministerio de Educación está emprendiendo una línea de acercamiento y trabajo conjunto con las organizaciones del sector educación y, entre ellas, con las organizaciones de la sociedad civil. Por otra parte, se han producido algunas prácticas innovadoras de trabajo con la sociedad a partir de datos, como puede ser la gymkana lanzada por la Secretaría de Educación de Bogotá para que el alumnado construya relatos digitales a partir de datos.

13. PRESENCIA DE UNA SOCIEDAD CIVIL ACTIVA EN EDUCACIÓN

En Colombia existe un grupo de organizaciones de la sociedad civil con una sólida trayectoria de colaboración para la mejora del sector educación, así como vías de financiación de esta actividad por parte del empresariado colombiano, como manera de contribuir a un mejor futuro del país.



ORGANIZACIONES DE LA SOCIEDAD CIVIL DESTACADAS:**ALIANZA EDUCACIÓN COMPROMISO DE TODOS**<http://www.educacioncompromisodetodos.org/>**DIVIDENDO POR COLOMBIA**<http://www.dividendoporcolombia.org/>**FUNDACIÓN COMPARTIR**<http://fundacioncompartir.org/>**FUNDACIÓN EXE (EMPRESARIOS POR LA EDUCACIÓN)**<http://fundacionexe.org.co/>**FUNDACIÓN CORONA**<http://www.fundacioncorona.org.co/>**MOVIMIENTO CIUDADANO TODOS POR LA EDUCACIÓN**<http://todosporlaeducacion.co/>

LA CAMPAÑA “EDUCACIÓN CINCO” ES UN EJEMPLO DE COLABORACIÓN ENTRE ENTIDADES DE LA SOCIEDAD CIVIL PARA CONTRIBUIR A MEJORAR LA EDUCACIÓN DEL PAÍS, EN 5 EJES:

1**ACCESO****2****PERMANENCIA****3****PERTINENCIA****4****CALIDAD****5****FINANCIACIÓN****Resumen de oportunidades detectadas**

#	OPORTUNIDAD	ÁREA
1	Diversidad de actores y complejidad del campo educación	1. Estrategia
2	Existencia de una política nacional de apertura de datos	1. Estrategia
3	Manejo de incentivos de los sistemas de medición	1. Estrategia
4	Buenas prácticas locales	3. Organización
5	Se han iniciado trabajos para una mejor relación con la información de empleo	3. Organización
6	Programas intensivos de evaluación de la calidad educativa	3. Organización
7	Renovación de sistemas de información	4. Tecnología
8	Existencia de aplicaciones a partir de datos	4. Tecnología
9	Sistemas para la mejora del reporte de datos	5. Datos
10	Muchos datos ya disponibles	5. Datos
11	Riqueza de datos estadísticos	5. Datos
12	Acercamiento a las organizaciones de la sociedad civil	6. Demanda
13	Presencia de una sociedad civil activa en educación	6. Demanda

REGIONALISMO CULTURA EMPRESA

PROYECTO PAIS

APROPIACION

ENTER.CO

HECHA LA LEY, HECHA LA TRAMPA

OPORTUNISMO

Puntos Vive Digital Plus

E

RESILIENCIA

Vive Gob Móvil

FAMILIA

CONFIANZA

CORRUPCIÓN

NERO

COLABORACION PARTICIPACION

PLANIFICACION

EXCLUSIÓN
ACCESO

D

BENEFICIOS DE UNA POLÍTICA DE APERTURA Y REUTILIZACIÓN DE DATOS

D.1. PARA QUÉ RETOS PUEDEN SER ÚTILES LOS DATOS ABIERTOS EN COLOMBIA

Colombia, tras años de esfuerzo en sus políticas educativas, es un país cuyas metas han superado los retos básicos de escolarización, para adentrarse en los de excelencia. Las dos grandes áreas donde se concentran los objetivos del Plan Nacional de Desarrollo son:

A AMPLIACIÓN DE COBERTURA CON CIERRE DE BRECHAS	B EXCELENCIA DOCENTE
JUNTO CON ESTAS ÁREAS, ENCONTRAMOS OBJETIVOS REGIONALES Y TAMBIÉN ESPECÍFICOS DE LA EDUCACIÓN SUPERIOR.	

En las entrevistas mantenidas, hemos comprobado que existe un alto grado de alineamiento en cuanto a cuáles son los principales retos a abordar ahora mismo. Algunos de estos retos tienen visos de poder beneficiarse de la apertura y uso de datos públicos. De manera muy general, serían los siguientes:

- 1 ENTENDER LA DESERCIÓN Y LANZAR MEDIDAS PARA CORREGIRLA
- 2 LIGAR COMPETENCIAS A LA EMPLEABILIDAD
- 3 EVALUAR IMPACTO DE POLÍTICAS DE EDUCACIÓN
- 4 CALIDAD EDUCATIVA, LIGADA A PROGRAMAS EDUCATIVOS
- 5 MEJORA DE LAS COMPETENCIAS DOCENTES
- 6 GARANTIZAR EL BUEN USO DE LOS RECURSOS PÚBLICOS EN LAS INSTITUCIONES EDUCATIVAS

LA LITERATURA EXISTENTE⁵ HABLA DE CINCO GRANDES GRUPOS DE BENEFICIOS DEL USO DE DATOS EN EDUCACIÓN, NOSOTROS HEMOS AÑADIDO UN SEXTO POR DELANTE, EL GRUPO 0, PARA EL CASO DE LATINOAMÉRICA:



0 - EDUCACIÓN PARA TODOS

Consiste en el compromiso para dar educación de calidad a todos los niños, jóvenes y adultos, con independencia de su lugar de residencia y de su nivel socioeconómico, logrando así una igualdad de oportunidades desde el inicio de la vida escolar. Los datos son fundamentales para entender la problemática, dar seguimiento a las medidas adoptadas e incluso para conseguir que las comunidades locales sean autónomas en la búsqueda de soluciones.



1 - MEJORA DE LA ENSEÑANZA

El uso de datos sobre el rendimiento de los alumnos, en relación con las estrategias didácticas, pueden ayudar a mejorar la manera en que se imparten las clases y a personalizar el aprendizaje. Para esto hay que analizar grandes cantidades de datos de individuos a lo largo de los años escolares, para obtener evidencia de cuáles son los métodos de enseñanza que mejor funcionan.



2 - ADECUACIÓN ENTRE ESTUDIANTES Y PROGRAMAS

Los datos pueden ayudar a los gestores públicos, pero también a padres y alumnos, al mejor encaje de las necesidades de cada alumnos en el programa y escuela más adecuados.



3 - ADECUACIÓN DE ESTUDIANTES Y EMPLEO

La finalidad es lograr el mejor encaje posible entre la demanda de empleo y las competencias adquiridas en el sistema educativo, potenciando la toma de decisiones informada por parte de padres y alumnos e implicando al sector empresarial en el análisis de las competencias relevantes.



4 - TRANSPARENCIA ECONÓMICO-FINANCIERA

El uso de datos sobre herramientas de análisis financiero permite entender el coste real de la educación y, una vez hechos transparentes los resultados, impulsa un mercado más eficiente y una mejor toma de decisiones en los casos en que es posible elegir.



5 - EFICIENCIA EN LA GESTIÓN

La disposición de datos comparativos del uso de recursos y de los resultados alcanzados puede mejorar sustancialmente la asignación de recursos y lograr mayores eficiencia y eficacia.

5 - McKinsey Global Institute (2013): Open data: Unlocking innovation and performance with liquid information.

D.2. POSIBLES BENEFICIOS PARA COLOMBIA DETECTADOS EN EL ESTUDIO DE CAMPO

A partir de las entrevistas, del análisis documental y de buenas prácticas existentes, se han determinado posibles elementos de valor a alcanzar mediante el uso de datos. Con ellos se ha elaborado una lista que no pretende ser exhaustiva, sino ofrecer una panorámica orientadora de posibles estrategias.

Se agrupan en cuatro grandes metas, con cierta relación con los planes nacionales de desarrollo y hemos señalado su relación con los 6 tipos de beneficios antes expuestos.

Este es el resultado:

METAS	ELEMENTOS DE VALOR A PARTIR DEL USO DE DATOS	0 educación para todos	1 mejor instrucción	2 adecuación estudiante programa	3 adecuación estudiantes - empleo	4 transparencia económica	5 eficiencia en la gestión
COBERTURA	Monitorizar el cumplimiento de una cobertura total	X					
	Estrategias de apoyo para la permanencia	X					
	Estrategias de apoyo para minimizar absentismo	X					
	Entender el desfase entre demanda y oferta	X					
	Caracterización socioeconómica de población estudiantil	X					
	Seguimiento de mejoras en infraestructuras rurales	X				X	
	Dar a conocer la planta de personal					X	X
	Dar a conocer la evolución de la matrícula			X	X	X	
	Dar a conocer la evolución de alumnos por docente		X			X	
	Dar a conocer metros cuadrados disponibles por nº de alumnos					X	
	Dar a conocer las jornadas disponibles		X			X	
	Orientar a la instalación de mayor capacidad donde se precise						X

METAS		0	1	2	3	4	5
ELEMENTOS DE VALOR A PARTIR DEL USO DE DATOS		educación para todos	mejor instrucción	adecuación estudiante - programa	adecuación estudiantes - empleo	transparencia económica	eficiencia en la gestión
CALIDAD EDUCATIVA	Mejor comprensión de las causas y efectos de la problemática educativa		X	X			X
	Respuesta rápida a malos resultados detectados		X	X			X
	Dar soporte a un sistema de calidad multidimensional						X
	Soporte a la gestión por resultados						X
	Información para toma de decisiones						X
	Análisis sistemático de la política educativa						X
	Integrar quejas de comunidad escolar en la gestión						X
	Rendición de cuentas y control social					X	
	Demostrar con datos que las políticas han sido (o no) útiles					X	
	Incentivo para mejorar la información					X	
	Análisis de programas educativos			X			
	Personalización de programas educativos			X	X		
	Seguimiento del programa de bilingüismo		X	X			
	Análisis de interacción en enseñanza media por computador		X				
	Establecimiento de incentivos a partir de mejores prácticas		X				
	Mejora de competencias docentes para mejores resultados		X				

METAS	ELEMENTOS DE VALOR A PARTIR DEL USO DE DATOS	0 <i>educación para todos</i>	1 <i>mejor instrucción</i>	2 <i>adecuación estudiante - programa</i>	3 <i>adecuación estudiantes - empleo</i>	4 <i>transparencia económica</i>	5 <i>eficiencia en la gestión</i>
GESTIÓN	Mejora de servicios prestados online a través de la analítica web	X		X			X
	Transparencia en la contratación					X	
	Transparencia en la gestión económica					X	
	Transparencia en la gestión de RRHH					X	
	Monitorizar nivel de ejecución presupuestal					X	X
	Conocer la inversión x alumno					X	
	Mejorar consistencia de información reportada					X	X
	Conocer la oportunidad de cubrir vacantes de docente					X	
	Conocer el costo por servicio					X	
	Mejora de procesos de atención ciudadana (p.e., convalidar títulos)						X
	Implicación comunidad educativa en gestión de recursos						X
	Implicación comunidad educativa en inspección de servicios						X
	Optimización del uso de recursos financieros						X

METAS	ELEMENTOS DE VALOR A PARTIR DEL USO DE DATOS	0 <i>educación para todos</i>	1 <i>mejor instrucción</i>	2 <i>adecuación estudiante - programa</i>	3 <i>adecuación estudiantes - empleo</i>	4 <i>transparencia económica</i>	5 <i>eficiencia en la gestión</i>
EQUIDAD Y DESARROLLO	Mejora del transporte escolar	X					
	Mejora de la nutrición escolar	X					
	Mejora de la salud escolar	X					
	Facilitar procesos burocráticos a los usuarios	X					
	Soporte a programas: embarazo, drogodependencia, malnutrición...	X					
	Seguimiento de "Ser Pilo Paga" y otros para aumentar igualdad	X					
	Seguimiento de enfoque de género	X		X	X		
	Poner en relación competencias profesionales y currículo educativo			X	X		
	Mejor orientación al empleo				X		
	Estudios de prospectiva laboral				X		







E

ANÁLISIS DE FUENTES DE DATOS PUBLICABLES

E.1. CATEGORIZACIÓN DE LAS FUENTES DE DATOS

El objetivo de este estudio es conocer la situación actual de los conjuntos de datos vinculados al sector educativo que se están publicando en el portal de datos abiertos de Colombia, compararla con la disponibilidad de datos abiertos similares en otras iniciativas internacionales relevantes, analizar el potencial de apertura a corto plazo de nuevos datasets relacionados con la cobertura y calidad educativa y analizar cómo se están utilizando y qué tipo de aplicaciones se están desarrollando con los datos abiertos en distintos países.

Para dar coherencia al estudio, y al objeto de establecer comparaciones sobre la tipología de datos que están publicando en sus iniciativas Open Data, tanto Colombia como otros países, es necesario tipificar el contenido de la información publicada en cada conjunto de datos, para además determinar qué tipo de datos son prioritarios de cara a optimizar su reutilización.

Cualquier categorización de datasets que se establezca debe responder a principios de utilidad de los datos abiertos. En el contexto educativo, los datos deben responder a usos que mejoren la calidad de la enseñanza, los estudiantes encuentren estudios a su medida, provoquen un encaje eficaz entre estudios, estudiantes y empleo, optimicen el uso de los recursos económicos incrementando la transparencia en su gestión así como lograr una administración lo más eficiente posible tanto en el momento presente como en su proyección futura.

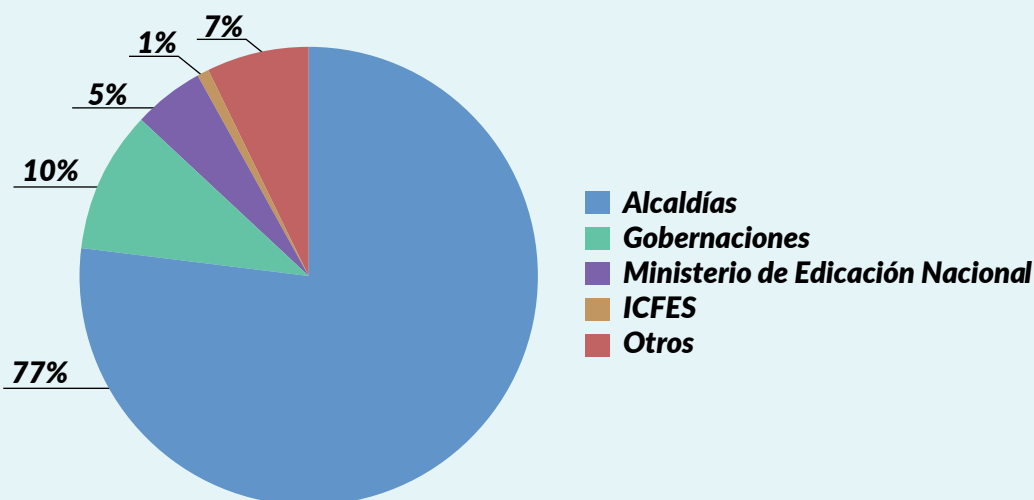
Partiendo de estas premisas, las categorías de datos que se proponen para facilitar la labor de estudio y comparación son las siguientes:

- 1** Datos relacionados con **INFRAESTRUCTURAS EDUCATIVAS**. Son aquellos datasets que se refieren a la identificación, descripción, localización y demás características de las instituciones y centros educativos.
- 2** Datos relacionados con **PROGRAMAS DOCENTES**. Agrupa los datasets que contienen información sobre programas educativos, especialidades ofertadas, calendarios docentes, niveles de formación, planes de estudios, actividades docentes y aspectos similares.
- 3** Datos **ESTADÍSTICOS**. Engloba información estadística de diversos tipos referentes a matriculaciones, estudios comparativos, proyecciones, encuestas, egresados, repitencias, deserciones, etc.
- 4** Datos **ECONÓMICOS**. Reflejan información sobre la gestión económica: gastos, inversiones, presupuestos, deuda y cuestiones similares.
- 5** Datos sobre **RECURSOS HUMANOS**. Se refiere a aquellos datos relativos con el equipo rectoral y docente, personal administrativo, gerentes de instituciones y recursos humanos en general.
- 6** Datos relativos a **SOPORTE INSTITUCIONAL**. Información relacionada con programas de apoyo, becas, dotaciones para equipamiento y demás.

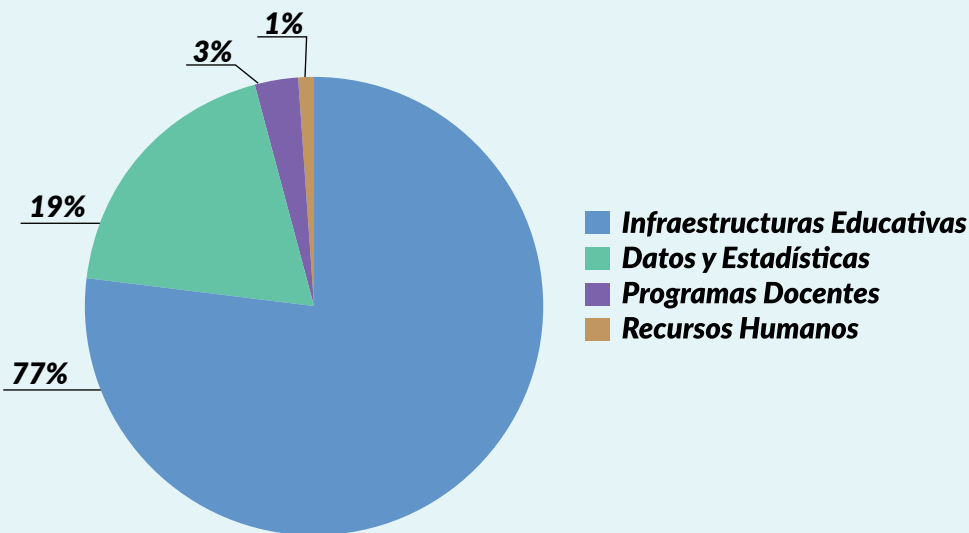
E.2. PUBLICACIÓN ACTUAL DE DATASETS EN EL PORTAL DE DATOS ABIERTOS DE COLOMBIA

A fecha de 29/10/2005, se publican en el portal de datos abiertos 221 conjuntos de datos relacionados con el sector educativo que se distribuyen de la siguiente forma según el tipo de entidad publicadora:

	Datasets
Alcaldías	171
Gobernaciones	22
MinEducación	10
ICFES	3
Otros	15



	Datasets
Infraestructuras Educativas	171
Datos y Estadísticas	41
Programas Docentes	6
Recursos Humanos	3



Del estado de publicación en el portal de datos abiertos se deduce:**1**

172 entidades distintas publican conjuntos de datos relacionados con el sector educación en el portal de datos abiertos. La mayoría de entidades (el 86%) publica un único dataset normalmente vinculado a la categoría de infraestructuras educativas. Las entidades que más datos publican son el Ministerio de Educación (10), la Gobernación de Huila (10), la Alcaldía de Tuluá (9), la Escuela Tecnológica Instituto Técnico Central (5), la Alcaldía de Tunja (5), el Instituto Nacional para Ciegos (4), la Alcaldía de Sogamoso (4), el Instituto Colombiano para el Fomento de la Educación Superior (3) y otras 16 entidades publican 2 conjuntos de datos.

2

Los datos que publican las distintas entidades del orden territorial (Alcaldías o Gobernaciones), tienen origen en los datos incorporados por cada Secretaría de Educación a SIMAT o de la investigación sobre la educación formal que publica el DANE referidos a cada uno de los territorios. A su vez, el Ministerio de Educación publica datos similares, cuyo origen fundamentalmente es SIMAT y los sistemas de información de la cobertura educativa, agregados por departamento, entidad territorial, municipio o consolidados del país. Esta situación produce una duplicidad de publicación de información en el portal de datos abiertos que puede producir inconsistencias de contenido.

3

La proporción de datasets relacionados con la cobertura educativa es muy superior a la oferta de datos relacionados con la calidad educativa. Tan sólo se publican 6 conjuntos de datos que contienen información sobre resultados de las pruebas SABER (3 de ellos del ICFES, los otros de la Gobernación de Huila y de la Alcaldía de Tutuá).

4

No se están publicando datasets de las categorías:
datos económicos y datos de soporte institucional.

El resultado de este análisis de datasets del portal de datos abiertos puede verse en el documento ***Comparativo Datasets_Educación_Internacional.xls***.

E.3. COMPARATIVO INTERNACIONAL DE PUBLICACIÓN DE DATOS ABIERTOS DEL ÁMBITO EDUCATIVO

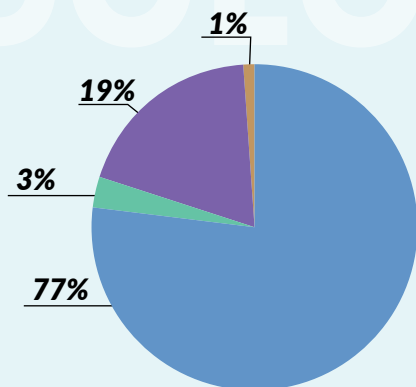
Complementario al análisis de datasets del ámbito educativo realizado sobre el portal de datos abiertos de Colombia, se realiza un estudio y tipificación de cada uno de los datasets publicados en las iniciativas de datos abiertos más relevantes, entre las que se encuentran las de México, Chile, Uruguay, Argentina, Ecuador, Brasil y España.

Hay que tener en cuenta que la forma de clasificar estos datasets no debe tomarse como única, ya que un mismo dataset puede encuadrarse en más de una categoría y en ocasiones no es evidente la tipificación de un dataset en función de los metadatos que lo describen. No obstante, y dado que el objetivo de la comparación no es tanto el obtener una contabilidad exhaustiva de los datasets publicados por cada iniciativa, sino destacar

aquellas iniciativas que presentan un mejor equilibrio en la publicación de conjuntos de datos cubriendo todas las categorías, la asignación ambigua de un tipo a un dataset no devalúa el objetivo del estudio.

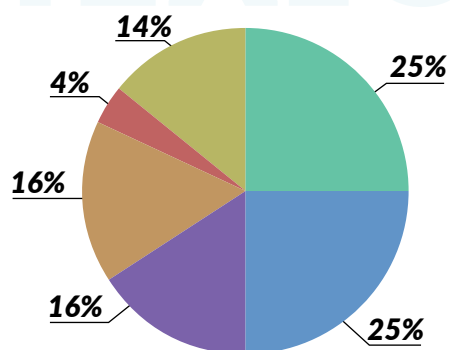
A continuación se detallan los resultados del estudio comparativo realizado en el periodo septiembre - octubre de 2015, sobre los conjuntos de datos del sector educativo publicados en las diferentes iniciativas seleccionadas. Los resultados de este análisis de iniciativas de datos abiertos puede verse en el documento Excel preparado ***Comparativo_Datasets_Educación_Internacional.xls***, en el que pueden encontrarse todos los datasets del ámbito educativo publicados por Colombia, México, Chile, Uruguay, Brasil, Argentina, Ecuador, España, USA, UK, Canadá e India.

COLOMBIA



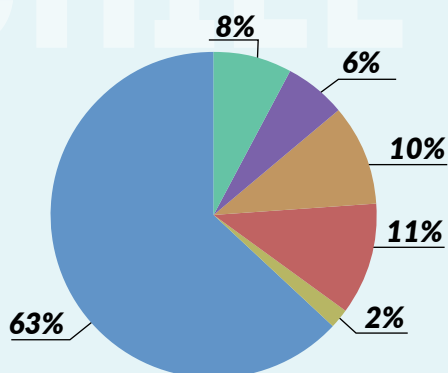
COLOMBIA		
	#	%
1- Infraestructuras educativas	171	77,38
2- Programas docentes	6	2,71
3- Datos y estadísticas docentes	41	18,55
4- Datos económicos	0	0,00
5- Recursos humanos	3	1,36
6- Soporte institucional	0	0,00
TOTAL	221	

MÉXICO



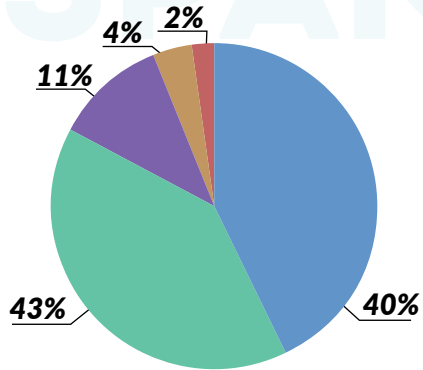
MÉXICO		
	#	%
1- Infraestructuras educativas	93	25,48
2- Programas docentes	58	15,89
3- Datos y estadísticas docentes	60	16,44
4- Datos económicos	13	3,56
5- Recursos humanos	51	13,97
6- Soporte institucional	90	24,66
TOTAL	365	

CHILE



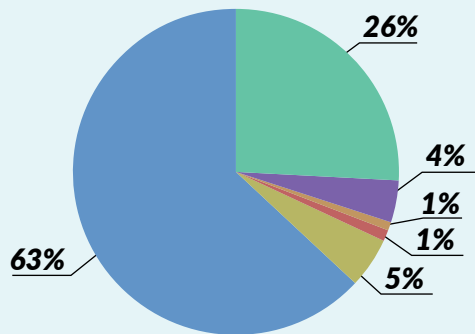
CHILE		
	#	%
1- Infraestructuras educativas	6	11,54
2- Programas docentes	1	1,92
3- Datos y estadísticas docentes	33	63,46
4- Datos económicos	4	7,69
5- Recursos humanos	3	5,77
6- Soporte institucional	5	9,62
TOTAL	52	

ESPAÑA



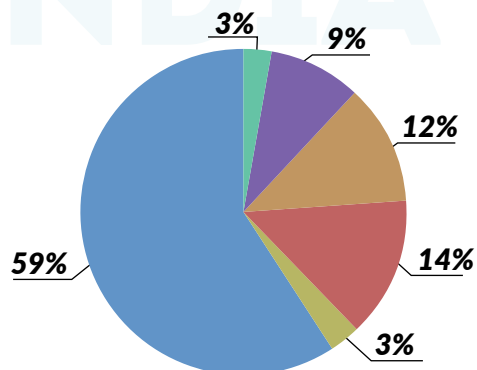
ESPAÑA	#	%
1- Infraestructuras educativas	76	11,55
2- Programas docentes	261	39,67
3- Datos y estadísticas docentes	281	42,71
4- Datos económicos	2	0,30
5- Recursos humanos	14	2,13
6- Soporte institucional	24	3,65
TOTAL	658	

CANADÁ



CANADÁ	#	%
1- Infraestructuras educativas	1	0,78
2- Programas docentes	7	5,43
3- Datos y estadísticas docentes	81	62,79
4- Datos económicos	34	26,36
5- Recursos humanos	5	3,88
6- Soporte institucional	1	0,78
TOTAL	129	

INDIA



INDIA	#	%
1- Infraestructuras educativas	32	14,55
2- Programas docentes	7	3,18
3- Datos y estadísticas docentes	129	58,64
4- Datos económicos	7	3,18
5- Recursos humanos	19	8,64
6- Soporte institucional	26	11,82
TOTAL	220	

Se ha excluido de la representación anterior aquellas iniciativas que no tienen un número relevante de datasets publicados, como es el caso de Uruguay (1 dataset), Brasil (20 datasets, donde el 85% son datos estadísticos), Argentina (1 dataset), Ecuador (1 dataset). Se excluyen también, por volumen, las iniciativas de UK (883 datasets) y USA (301 datasets). No obstante, en el documento Excel preparado se detallan toda la información de las iniciativas analizadas.

Como conclusión principal de este análisis comparativo se deduce que el país con un mejor planteamiento de apertura de datos del sector educativo es México, ya que por un lado publica datasets de todas las categorías de información y además presenta el mejor equilibrio en cuanto a cantidad de datasets publicados, de todas las iniciativas analizadas.

E.4. CONJUNTOS DE DATOS POTENCIALMENTE PUBLICABLES COMO OPEN DATA

DATOS A PARTIR DE LOS SITIOS WEB DE EDUCACIÓN

En el desarrollo de este estudio se ha llevado a cabo un análisis de potenciales datasets publicables en el portal de datos abiertos.

Para su realización se han seguido los siguientes criterios:

- Localizar archivos que contienen datos y que actualmente están publicados en algún portal Web del ámbito educativo.
- Identificar aquellos archivos de datos en formatos estructurados (XLS, CSV, etc.) que están disponibles para su descarga sin restricciones de acceso.
- Identificar paginas o aplicaciones web que presentan datos no descargables pero de forma estructurada en formato HTML y que por su potencial de reutilización merece la pena su conversión en archivo descargable.
- Seleccionar del catálogo de sistemas de información de MinEducación aquellos que procesan información susceptible de generar datos abiertos.

Este estudio debe tomarse como base para un proceso de identificación más amplio, ya que si bien se ha localizado una cantidad relevante de potenciales datasets, existen muchas aplicaciones web y sistemas que requieren acceso de usuario registrado para realizar el proceso de identificación. Por otra parte, es necesario además realizar un estudio del contenido de los diversos archivos descargables en formato Excel al objeto de verificar que cumplen criterios de calidad para su reutilización.

El resultado del estudio arroja la identificación de conjuntos de datos en los siguientes sitios web o aplicaciones:

- Datos en la aplicación BI (Business Intelligence) de Educación preescolar, básica y media
- Datos en la aplicación BI (Business Intelligence) de Educación Superior
- Datos en el Sistema de Información de la Educación para el Trabajo y el Desarrollo Humano (SIET)
- Datos estadísticos del Sector Educativo
- Datos en el Sistema de Seguimiento (sistema con el que se hace seguimiento al reporte de la planta docente y la matrícula de educación básica y media)
- Datos en el Sistema para la prevención de la deserción de la Educación Superior (SPADIES)
- Datos relacionados con la calidad educativa publicados en el servicio FTP de ICFES.
- Datos en el Observatorio Laboral para la Educación (OLE)
- Datos relacionados con la gestión financiera del MinEducación
- Datos relacionados con la contratación de bienes y servicios del MinEducación
- Datos relacionados con el sector educativo que publica el DANE
- Datos del catálogo de recursos educativos
- Datos en Modelo de Indicadores del Desempeño de la Educación (MIDE)
- Datos existentes en otros sistemas de información o aplicaciones Web que sirven datos para su descarga pública
- Datos en la aplicación Buscando Colegio
- Datos en la aplicación Buscando Carrera
- Datos relacionados con Educación publicados en el portal de datos abiertos de Bogotá
- Datos relacionados con Educación publicados en el portal de datos abiertos de Tuluá
- Datos relacionados con Educación publicados en el portal de datos abiertos de ESRI

Todos los datos identificados están listados en el documento Excel preparado de apoyo a este estudio: **Identificacion_Datasets_MinEducacion.xlsx.**

Se han identificado más de 600 archivos con datos que son potenciales datasets publicables a corto plazo como datos abiertos ya que éstos están actualmente disponibles para su descarga en formato XLS o CSV principalmente.

Esta cantidad de datasets puede ser muy superior ya que en muchos de los sistemas habilitados para el acceso y descarga de estos ficheros de datos permiten parametrizar la configuración del contenido en función de la selección por parte del usuario de diferentes variables.

Entre otras cuestiones se debe tener en cuenta:

- Muchas de las aplicaciones web que permiten descarga de datos se pueden parametrizar por su cobertura territorial pudiendo seleccionar el ámbito nacional o los distintos departamentos.
- Algunas consultas permiten la selección por tipo de población estudiantil (matriculados, admitidos, inscritos, primer curso), metodologías de programa (a distancia tradicional, a distancia virtual o presencial), sector de la institución educativa (oficial o privada) entre otros filtros de selección.

- En el caso de los resultados de calidad educativa de ICFES tan sólo se han identificado los datos más actuales correspondientes a las anualidades 2013, 2014, pero se encuentran datos históricos desde 2006.
- Otros filtros que se pueden aplicar se basan en la selección de consolidados nacionales, por secretarías o por establecimientos educativos.

Por otro lado, hay que tener en cuenta la posible duplicidad de conjuntos de datos actualmente publicada en las distintas secciones web de los portales de Educación, como es el caso por ejemplo, de los datos ofrecidos en la sección de estadísticas del sector educativo y los ofrecidos a través del sistema de seguimiento. Por tanto es necesario verificar que los archivos de datos descargables en ambas secciones no se solapan.

Una estimación factible una vez verificada la no duplicidad de datasets publicables y la selección de variables adecuadas para la aplicación de filtros nos lleva a pensar que la cifra de potenciales datasets se sitúe en torno a los 500 conjuntos de datos.

En cuanto a las categorías de información contenida en los conjuntos de datos identificados se obtiene:

DATASETS	CATEGORÍAS DE INFORMACIÓN			
	Infraestructuras educativas	Datos estadísticos	Recursos humanos	Soporte institucional
BI_BASICA Y MEDIA	Infraestructuras educativas	Datos estadísticos		
BI_SNIES	Datos estadísticos	Recursos humanos		
BI - SIET	Infraestructuras educativas	Programas docentes	Datos estadísticos	
Estadísticas Sector Educativo	Datos estadísticos	Datos económicos	Recursos humanos	Soporte institucional
Sistema de Seguimiento	Datos estadísticos	Recursos humanos	Soporte institucional	
SPADIES	Datos estadísticos	Soporte institucional		
ICFES	Datos estadísticos			
OLE	Datos estadísticos	Datos económicos	Programas docentes	
Gestión Financiera	Datos económicos			
Contratación MEN	Datos económicos			
DANE	Datos estadísticos			
Recursos Educativos	Programas docentes			
MIDE	Datos estadísticos			
OTROS	Varios			
Buscando Colegio	Infraestructuras educativas			
Buscando Carrera	Infraestructuras educativas	Programas docentes		
OD Bogotá	Infraestructuras educativas	Datos estadísticos	Datos económicos	
OD Tuluá	Infraestructuras educativas	Datos estadísticos		
OD ESRI	Datos estadísticos			

De la anterior categorización de contenidos se deriva que se cubren la totalidad de las categorías de datos para equiparar la iniciativa de datos abiertos del sector educativo con países que presentan una mejor situación de equilibrio como es el caso de México.

Se detectan además las siguientes carencias en la información publicada:

- A pesar de que se publican datos relacionados con infraestructuras educativas de las diferentes instituciones, en lo referente a datos sobre dotaciones para la enseñanza de los centros educativos como son aulas de informática, laboratorios, computadores, tabletas, acceso a internet, disponibilidad wifi, antigüedad de equipamientos, obras de adecuación, etc., tan solo se publican, en el sistema de seguimiento de planta docente y matrícula de educación básica y media, datos de alumnos por computador, dotación de PC's, dotación de conectividad e indicadores de conectividad. En el ámbito de la educación superior, no se publica nada al respecto.
- No se publican datos económicos a nivel de secretarías de educación relacionados con gasto o inversión en los centros educativos aunque los datos deben estar en SIIF (aplicación transversal a todas las entidades del estado gestionada por el Ministerio de Hacienda y Crédito Público para el control de información económico financiera). El portal de transparencia económica que expone públicamente datos extraídos de SIIF muestra esta información a nivel de entidad tan sólo para algunas entidades.
- Los datos relacionados con los recursos asignados a cada institución educativa están parcialmente disponibles a través del sistema CHIP (Consolidador de Hacienda e Información Pública (CHIP). Se recogen en este sistema datos del PAE (Programa de Alimentación Escolar). Este sistema permite a las entidades territoriales hacer un reporte trimestral de la gestión financiera del PAE, para que el Ministerio de Educación realice el seguimiento y monitoreo a los recursos destinados a la Alimentación Escolar en los establecimientos educativos oficiales en el país. Por su parte, el Ministerio de Educación que conoce los datos sobre el suministro de raciones alimenticias a cada centro educativo por los contratos de servicios que gestiona tampoco publica estos datos.

Por otro lado, aunque existen muchos datos descargables desde diferentes secciones de la web de MinEducación es importante señalar que **la política de privacidad y condiciones de uso⁶ del portal del Ministerio de Educación no es plenamente compatible con la vocación de reutilización** de la información pública promovida por la iniciativa de datos abiertos ya que restringe la reutilización con fines comerciales. En dicha política se encuentra el siguiente párrafo:

“Se prohíbe cualquier uso, transformación o explotación de los contenidos incluidos en el Sitio Web con finalidades comerciales o promocionales salvo autorización previa del Ministerio de Educación; en cualquier caso se prohíbe cualquier uso contrario a la ley y del Sitio Web para uso personal y no comercial, siempre y cuando se haga expresa mención de la propiedad en cabeza del autor del contenido.”

Por otro lado, no se está asegurando una descarga segura de archivos de datos lo que puede mermar el potencial de la reutilización.

“El Ministerio de Educación no controla ni garantiza la ausencia de virus ni de otros elementos en los contenidos que puedan producir alteraciones en su sistema informático (software y hardware) o en los documentos electrónicos y ficheros almacenados en su sistema informático.”

Por el contrario y relacionado con este respecto, en los términos de uso⁷ del sitio web “Colombia Aprende” se indica:

“Respecto de aquellos programas, materiales o información que se haya puesto a disposición del usuario para que éste pueda descargarlos en su equipo, sistema o cualquier medio o soporte de memoria de su propiedad o cuya tenencia detente (“materiales de download”) por parte de Colombia Aprende o cualquier proveedor de información, Colombia Aprende informa que inicialmente dichos programas fueron revisados a efectos de considerarlos libres de virus, u otros medios tecnológicos de tipo destructivo.”

Por tanto, si los datasets identificados como potenciales conjuntos de datos abiertos se publican como tal será necesario revisar las condiciones de uso de los sitios web vinculados al Ministerio de Educación.

6 - Políticas de privacidad y condiciones de uso del portal web del MEN: <http://www.mineducacion.gov.co/1621/w3-article-336187.html>

7 - Términos y condiciones de uso del portal web “Colombia Aprende”: http://www.colombiaaprende.edu.co/html/home/1592/article-68350.html#h2_6

DATOS A PARTIR DE LAS APLICACIONES VERTICALES

Una segunda parte de este estudio sobre conjuntos de datos potencialmente publicables como Open Data es la referida a la aportación que pueden realizar las aplicaciones verticales de gestión disponibles en el ámbito del Ministerio de Educación. Estas aplicaciones son el origen de los datos que se publican en los portales Web del Ministerio pero además pueden ser la fuente para la apertura de nuevos datasets.

MinEducación dispone de 115 sistemas de información, según la información facilitada por la Oficina de Tecnologías y Sistemas de Información. Estos sistemas cubren los ámbitos de educación básica, educación superior, planeación, Ministerio y otros de carácter interno. Para realizar este estudio se dispone únicamente de una descripción de la funcionalidad del sistema, insuficiente para afirmar categóricamente qué sistemas son susceptibles de generar datos abiertos.

No obstante, y en base a la información disponible, se realiza una distinción entre los sistemas que pueden contener datos abiertos, los que requieren una revisión y posterior toma de decisión en aquellos cuya descripción genera dudas y los que a priori no procesan datos susceptibles de apertura o son de carácter interno. De este somero análisis se obtiene que existen al menos 38 sistemas de información que podrían ser orígenes de datos abiertos.



ÁMBITO	SUSCEPTIBLES DE APERTURA	REVISABLES	NO SUSCEPTIBLES DE APERTURA	TOTALES
Educación básica y media	21	3	13	37
Educación superior	12	2	14	28
De uso interno	1	4	25	30
Del Ministerio de Educación	1	0	6	7
Planeación	1	0	0	1
Sin asignación de ámbito	2	1	9	12
	38	10	67	115

La información detallada de estos sistemas de información se encuentran en el documento Excel preparado ***Inventario_Sistemas_Información_2015_MinEducacion.xlsx***.



Términos y condiciones de uso del portal web "Colombia Aprende":
http://www.colombiaaprende.edu.co/html/home/1592/article-68350.html#h2_6

F

ESTUDIO DE LA DEMANDA

F.1. DEMANDA INTERNA DE DATOS DE EDUCACIÓN

El estudio de la demanda es importante ya que permite determinar las necesidades de datos que se producen entre entidades constituyendo un espacio fundamental para la toma de decisiones y priorización de apertura de datasets.

Desde el punto de vista de la demanda interna de datos al propio Ministerio y las entidades bajo su influencia, cabe mencionar los flujos de información señalados en el **Documento de diseño y especificación de la arquitectura. Solución de capacidades de Interoperabilidad - Vista de información** del documento del marco de referencia de Arquitectura Empresarial, en el que se identifica la información más significativa que intercambian las entidades cabeza de sector, con otras entidades y otros organismos privados o internacionales, es decir, los datos de entrada y salida o lo que es lo mismo, entidad que envía y entidad que recibe datos.

Los flujos de información se realizan fundamentalmente en base a acuerdos de intercambio de información o convenios de colaboración para el intercambio de información realizados ad hoc, es decir, no existe un marco común de interoperabilidad. Desde el punto de vista técnico, apenas se han hecho disponibles una serie

de Web Services con algunas entidades como ICFES o DANE. Normalmente, los intercambios se realizan mediante envío de bases de datos o disponibilidad de servicios FTP.

Con el objeto de paliar esta situación se desarrollará el sistema RENE (Registro Nacional de Educación), que pretende levantar la arquitectura empresarial del sector (bus de integración) y constituirse en estándar para el intercambio de información. Es un proyecto a largo plazo; actualmente en fase de especificación y conceptualización (2015 - 2016) y su implementación está prevista para 2018.

A continuación se presenta la entrada y salida de información del MinEducación.



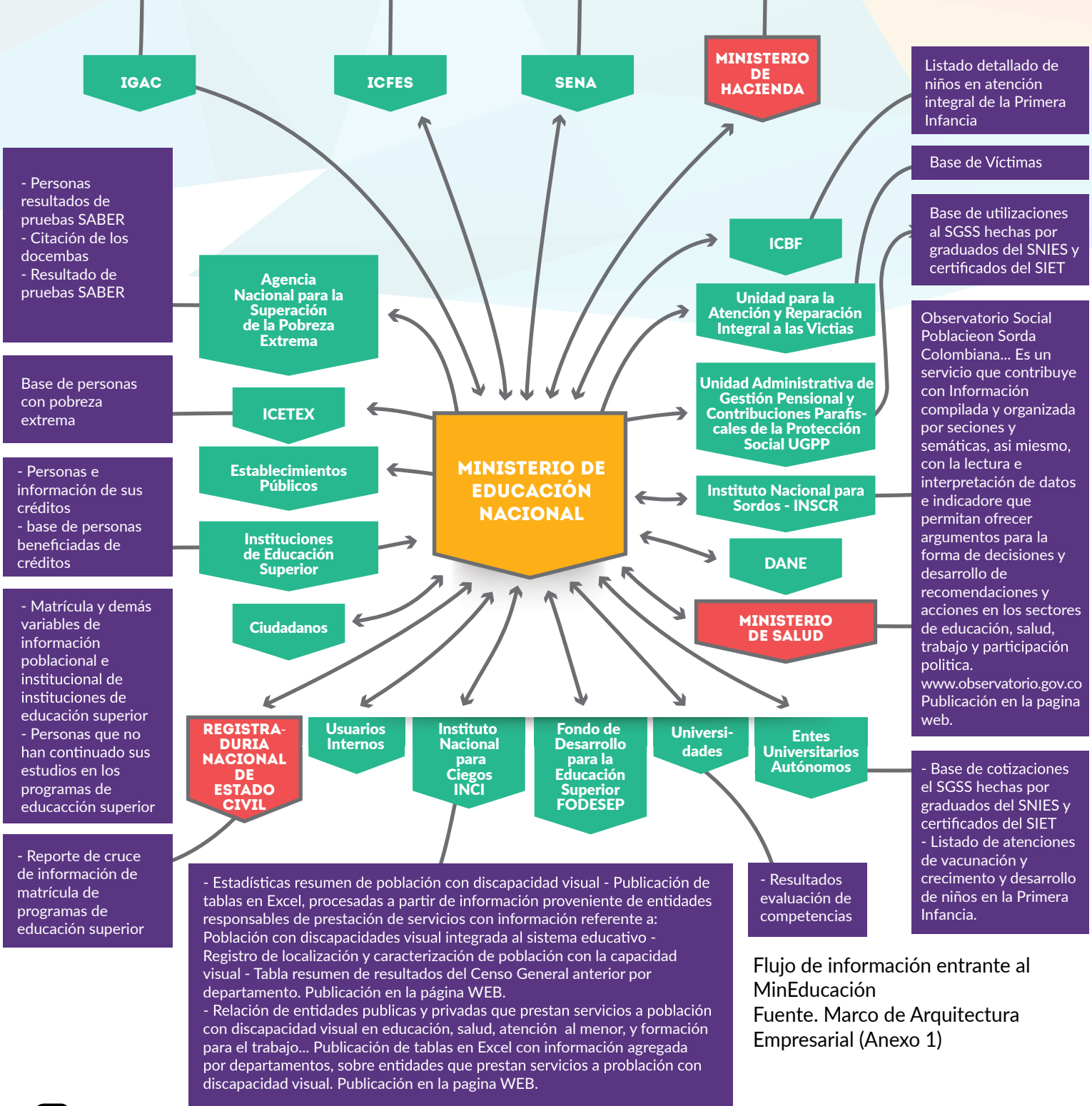
ENTRADA DE INFORMACIÓN:

Intercambio de datos para el Sistema de Información Geográfico para la planeación y el ordenamiento territorial - SIGOT. Afianza la cooperación, integración e intercambio de datos y servicios entre las entidades participantes en el Sistema de Información Geográfica para la Planeación y el Ordenamiento Territorial en Colombia SIG-OT, para proveer información y servicios a las autoridades y estancias del Sistema de Planeación en Colombia, a nivel nacional, departamental y municipal. Servicio WEB.

- Personas resultados de pruebas SABER
- Citación de los docembas
- Resultado de pruebas SABER

Personas matriculadas y certificadas el SENA

RP, COP, Pegas, Solicitud de Vigencias Futuras y Compromiso de vigencias futuras



Flujo de información entrante al MinEducación
Fuente. Marco de Arquitectura Empresarial (Anexo 1)

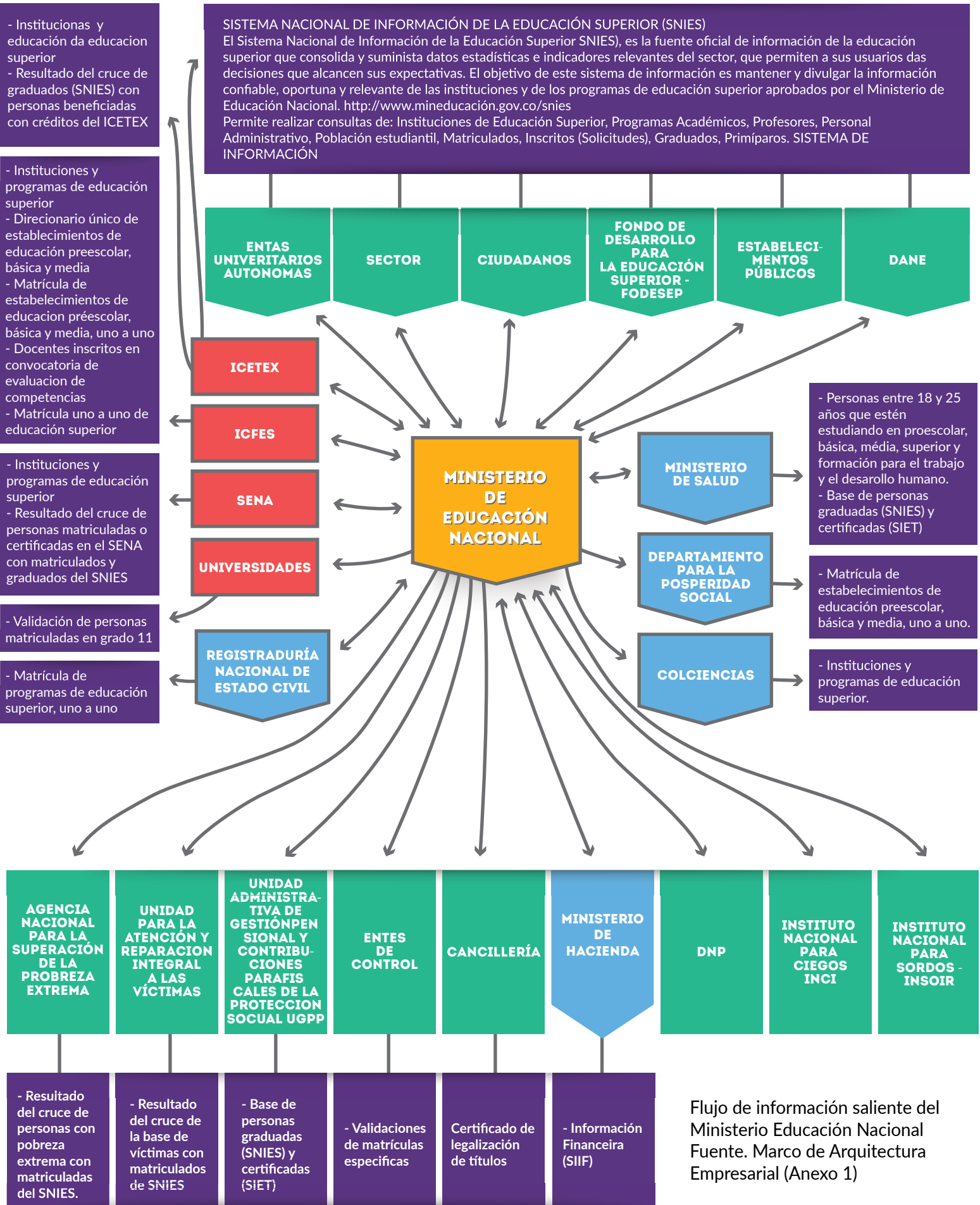
El Ministerio recibe la siguiente información:

ACTOR ENTIDAD QUE ENVÍA INFORMACIÓN	INFORMACIÓN QUE RECIBE⁸
SENA.	Personas matriculadas o certificadas en el SENA.
Ministerio de Hacienda.	RP, CDP, Pagos, solicitud de vigencias futuras y compromiso de vigencias futuras.
Instituto Colombiano de Bienestar Familiar – ICBF.	Listado detallado de niños en atención integral de la primera infancia.
U. para la Atención y Reparación Integral a las Víctimas.	Base de víctimas.
U. Administrativa de Gestión Pensional y Contribuciones Parafiscales de la Protección Social UGPP.	Base de cotizaciones al SGSS hechas por graduados del SNIES y certificados del SIET.
Instituto Nacional para Sordos – INSOR.	Información compilada y organizada por sectores y temáticas orientadas a inclusión social.
Departamento Administrativo Nacional de Estadística – DANE	*
Ministerio de Salud.	Base de cotizaciones al SGSS hechas por graduados del SNIES y certificados del SIET.
	Listado de atenciones de vacunación y crecimiento y desarrollo de niños en la primera infancia.
Entes universitarios autónomos.	Resultados evaluación de competencias.
Centros de Educación Superior.	Resultados evaluación de competencias.
Fondo De Desarrollo Para La Educación Superior – FO-DESEP.	Estadísticas resumen de población con discapacidad visual.
Instituto Nacional para Ciegos –INCI.	Relación de entidades públicas y privadas que prestan servicios a población con discapacidad visual en educación, salud, atención al menor, y formación para el trabajo.
Usuarios internos.	*
Registraduría Nacional de Estado Civil.	Reporte de cruce de información de matrícula de programas de educación superior.
Ciudadanos.	*
Instituciones de Educación Superior.	Matricula y demás variables de información poblacional e institucional de instituciones de educación superior.
	Personas que no han continuado sus estudios en los programas de educación superior.
Establecimientos públicos.	*
Instituto Colombiano de Estudios Técnicos en el Exterior – ICETEX.	Personas e información de sus créditos.
	Base de personas beneficiadas de créditos.
Agencia Nacional para la Superación de la Pobreza Extrema.	Base de personas con pobreza extrema.
Instituto Geográfico Agustín Codazzi – IGAC.	Intercambio de datos para el Sistema de Información Geográfico para la planeación y el ordenamiento territorial.
Instituto Colombiano para la Evaluación de la Educación – ICFES.	Personas y resultados de pruebas SABER.
	Citación de los docentes.
	Resultado de pruebas SABER (uno a uno).

Entradas de información del MinEducación

8 - Aquellos campos que tengan un * hacen referencia a los campos con los cuales no se contó con información.

SALIDA DE INFORMACIÓN:



El Ministerio entrega información a los siguientes actores:

ACTOR / ENTIDAD QUE RECIBE INFORMACIÓN	INFORMACIÓN QUE RECIBE ⁹
Instituto Colombiano de Estudios Técnicos en el Exterior – ICETEX.	Estadísticas e indicadores relevantes del sector.
Entidades del sector.	
Ciudadanos.	
Fondo De Desarrollo Para La Educación Superior – FODESEP.	
Establecimientos Públicos.	
Departamento Administrativo Nacional de Estadística – DANE.	
Ministerio de Salud.	Personas entre 18 y 25 años que estén estudiando en preescolar, básica, media, superior y formación para el trabajo y el desarrollo humano.
	Base de personas graduadas del Sistema Nacional de Información de la Educación superior (SNIES) y certificadas del Sistema de Información de Educación para el Trabajo y Desarrollo Humano (SIET).
Departamento para la Prosperidad Social.	Matricula de establecimientos de educación preescolar, básica y media, uno a uno.
Departamento Administrativo de Ciencia, Tecnología e Innovación – COLCIENCIAS.	Instituciones y programas de educación superior.
Instituto Nacional para Sordos – INSOR.	*
Instituto Nacional para Ciegos – INCI.	*
Departamento Nacional de Planeación – DNP.	*
Ministerio de Hacienda.	Información Financiera – SIIF.
Cancillería.	Certificado de legalización de títulos.
Entes de control.	Validaciones de matrículas específicas.
U. Administrativa de Gestión Pensional y Contribuciones Parafiscales de la Protección Social UGPP.	Base de personas graduadas (SNIES) y certificadas (SIET).
Unidad para la Atención y Reparación Integral a las Víctimas.	Resultado del cruce de la base de víctimas con matriculados del SNIES.
Registraduría Nacional de Estado Civil.	Matricula de programas de educación superior, uno a uno.
Agencia Nacional para la Superación de la Pobreza Extrema.	Resultado del cruce de personas con pobreza extrema con matriculados del SNIES.
Centros de Educación Superior.	Validación de persona matriculada en grado 11.
Servicio Nacional de Aprendizaje – SENA.	Instituciones y programas de educación superior.
	Resultado del cruce de personas matriculadas o certificadas en el SENA con matriculados y graduados del SNIES.
Instituto Colombiano para la Evaluación de la Educación – ICFES.	Instituciones y programas de educación superior.
	Directorio único de establecimientos de educación preescolar, básica y media.
	Matricula de establecimientos de educación preescolar, básica y media, uno a uno.
	Docentes inscritos en convocatoria de evaluación de competencias.
Instituto Colombiano de Estudios Técnicos en el Exterior – ICETEX.	Matrícula uno a uno de educación superior.
	Instituciones y programas de educación superior.
	Resultado del cruce de graduados (SNIES) con personas beneficiadas con créditos del Instituto Colombiano de Crédito Educativo y Estudios en el Exterior (ICETEX).

Salidas de información del MinEducación

⁹ - Aquellos campos que tengan un * hacen referencia a los campos con los cuales no se contó con información.



F.2. DEMANDA EXTERNA DE DATOS DE EDUCACIÓN

Desde el punto de vista de la demanda externa de información, además de constituir un elemento fundamental del ecosistema de datos abiertos es clave por el ejercicio de retroalimentación que realizan por medio de publicaciones, tanto de informes como de datos.

La alianza Educación Compromiso de Todos¹⁰, a la que pertenecen La Fundación Empresarios por la Educación, la Fundación Saldarriaga Concha, la Fundación Corona, la Fundación Restrepo Barco, Unicef Colombia, la Universidad de los Andes y la Corporación Región, tiene como misión promover la garantía integral del derecho a la educación en Colombia. Desarrollan el Índice del Derecho a la Educación – IDE¹¹, que es una herramienta de consulta que permite conocer cómo va Colombia en el cumplimiento de la obligación del derecho a la educación y cómo están las 94 entidades territoriales certificadas teniendo en cuenta un promedio nacional. El índice se creó a partir de 62 variables distribuidas en cuatro dimensiones: Disponibilidad, Accesibilidad, Adaptabilidad y Aceptabilidad en base a datos de matrícula, ruralidad o discapacidad obtenidos fundamentalmente del MinEducación y del INFES con mucha dificultad al no ser datos abiertos, según declaran los responsables del índice. Los datos que contiene están en formato Excel descargable y corresponden a 2012 y 2013.

La Fundación Dividendo por Colombia¹², tiene como misión promover la conciencia y la práctica de la responsabilidad social empresarial e individual y canalizarla hacia la mejora de la calidad de vida de las comunidades marginadas de Colombia. Entre sus actividades desarrolla programas educativos en alianza público privada para elevar el indicador de la permanencia de los niños y jóvenes en el colegio. Para la realización de esta tarea requieren de datos de matrícula del MEN que no están disponibles con el nivel de desagregación que requieren, ya que el enfoque de sus programas tienen una orientación hacia la prevención de la deserción escolar y para ello pretenden elaborar modelos predictivos para construir nuevos indicadores a partir de la información de matrícula en los centros escolares.

La Fundación Compartir¹³, es una fundación empresarial que contribuye a mejorar la calidad de la educación de niños y jóvenes colombianos. Entre otros trabajos de investigación han desarrollado el estudio “Tras la excelencia docente” en el que se plantea que la calidad de la educación depende de la calidad de la formación de los docentes. La realización de este tipo de trabajos de investigación implica el consumo de datos de cobertura, calidad e infraestructura educativa además de la información completa y actualizada sobre la plantilla docente. La demanda fundamental de esta organización, además de la disponibilidad de los datos es la necesidad de mejorar la encontrabilidad de la información.

10 - Alianza Educación compromiso de todos: <http://www.educacioncompromisodetodos.org/>

11 - Índice del derecho a la Educación – IDE: <http://www.educacioncompromisodetodos.org/derecho-a-la-educacion/>

12 - Fundación Dividendo por Colombia: <http://www.dividendoporcolombia.org/es/>

13 - Fundación Compartir: <http://fundacioncompartir.org/>

Por otro lado, cabe reseñar el caso de la Fundación Corona patrocinadora de la red “Como Vamos” que ejerce un doble rol como consumidor / productor de datos abiertos.

En su rol como productor de datos están lanzando una iniciativa de Open Data¹⁴, en colaboración con EOKOE de Brasil y CODEANDO de México, con tres funcionalidades: plataforma de datos abiertos de la Red de Ciudades Como Vamos, actualmente 14 ciudades (implementada sobre CKAN), la Aplicación de datos abiertos (en desarrollo, ofrecerá visualizaciones de datos) y la funcionalidad de Storytelling (sobre Wordpress, para que académicos y ciudadanos activos puedan contar historias utilizando los datos de la plataforma).

Como consumidores de datos utilizan los siguientes Indicadores recopilados a nivel territorial en cada ciudad de las que integran la red “Como Vamos”.

- Cobertura neta en educación inicial (prejardín y jardín)
- Tasa de cobertura neta en preescolar
- Tasa de cobertura neta en Primaria
- Tasa de cobertura neta en Secundaria
- Tasa de cobertura neta en Media
- Tasa de asistencia escolar de 5 a 16 años
- Tasa de deserción escolar - sector público
- Porcentaje de estudiantes de 5º con un nivel avanzado o satisfactorio en las pruebas estandarizadas nacionales de matemáticas
- Porcentaje de estudiantes de 5º con un nivel avanzado o satisfactorio en las pruebas estandarizadas nacionales de lenguaje
- Porcentaje de estudiantes de 9º con un nivel avanzado o satisfactorio en las pruebas estandarizadas nacionales de matemáticas
- Porcentaje de estudiantes de 9º con un nivel avanzado o satisfactorio en las pruebas estandarizadas nacionales de lenguaje
- Porcentaje de estudiantes de 11º con dominio del inglés con nivel B1 o B+

Su necesidad principal es localizar datos para estos indicadores desagregados por localidad o comuna, género y edad. Otro dato que demandan desde esta Fundación es el porcentaje del presupuesto público dedicado a la educación por ciudad.

Estos datos se utilizan para construir el Índice del Derecho a la Educación que se basa en 4 grupos de indicadores: disponibilidad, accesibilidad, adaptabilidad y aceptabilidad. En el documento Excel realizado bajo este estudio **Indicadores_Fundacion_Corona.xls** se recoge la información sobre esta iniciativa.



14 - Portal Open Data “Como Vamos”: <http://comovamos.eokoe.com/organization>

G

RECOMENDACIONES PARA LA ACCIÓN

A continuación, ofrecemos una serie de recomendaciones encaminadas a guiar un plan de acción. El apartado de datos se ofrece por separado, por su mayor grado de detalle.

G.1.

RECOMENDACIONES GENERALES

	ÁMBITO	RECOMENDACIONES
1	1. Liderazgo	Impulsar la publicación de datos abiertos en el portal nacional
2	1. Liderazgo	Dar los pasos para lograr un currículo educativo a lo largo de toda la vida del individuo, soportado en un identificador único que permita la anonimización de los registros
3	2. Legal	Elaborar normativa interna que impulse la publicación de datos por defecto, siguiendo la Carta de los Principios Internacionales para los Datos Abiertos ¹⁵
4	2. Legal	Realizar un análisis completo de los casos en que debe anteponerse el habeas data a la publicación de datos, para evitar la necesidad de revisar caso a caso
5	2. Legal	Revisar las condiciones de uso de los portales de MinEducación
6	3. Organización	Sistemas de gestión del conocimiento para evidenciar y emular la innovación local de los establecimientos educativos y las secretarías de educación
7	3. Organización	Estrechar la colaboración con SENA y con el servicio nacional de empleo para compartir información a todo lo largo de la vida de los individuos
8	3. Organización	Delimitar la información publicable por las Alcaldías y Gobernaciones para evitar duplicidades con la publicación de datos del Ministerio y entidades del orden Nacional
8A	3. Organización	Entidades territoriales podrían publicar (información de competencia exclusivamente local): plazas disponibles, becas y subsidios educativos, beneficiarios de alimentación escolar, obras y reparaciones en instituciones educativas, planes de mejoramiento, ejecuciones presupuestarias, incentivos para docentes, convocatorias de créditos educativos, licencias educativas (privadas), etc
8B	3. Organización	MinEducación publicaría información consolidada relativa a infraestructura educativa (incluida aquella que reportan las entidades territoriales una vez validada): datos de matrícula, programas docentes, resultados de calidad educativa, datos estadísticos y aquellos datos que posee con garantía de calidad y actualización
9	3. Organización	Mantener los esfuerzos actuales para mejora de la calidad del datos y publicar reportes acerca de los resultados del programa
10	3. Organización	Detectar y difundir mejores prácticas locales en la creación de un ecosistema para los datos abiertos, como es el caso de Bogotá
11	3. Organización	Implantar sistemas de incentivos para quienes producen buenos datos
12	4. Tecnología	Aprovechar como oportunidad el futuro portal de calidad educativa de ICFES para publicar datos abiertos y para federarlos con el portal nacional
13	4. Tecnología	Interoperabilidad: valorar la posibilidad de que datos abiertos sean la fuente única de intercambio de información entre entidades
14	4. Tecnología	Interoperabilidad: aprovechar la disponibilidad de APIs en la nueva plataforma de datos abiertos para mejorar la eficiencia del marco de interoperabilidad entre entidades
15	6. Demanda	Desarrollar un plan de comunicación de datos abiertos para educación con mensajes adaptados a los diferentes actores/reutilizadores de datos
16	6. Demanda	Incrementar las oportunidades de participación de cada actor en el sector educación
17	6. Demanda	Establecer un diálogo continuo con el sector empresarial para que comprenda los problemas del sector educación y pueda aportar soluciones
18	6. Demanda	Facilitar un ecosistema de reutilización mediante eventos, talleres y otras actividades de fomento de la reutilización, donde se sociabilicen los retos y metas en el sector con la comunidad reutilizadora
19	6. Demanda	Lanzar campañas para implicar a la sociedad en la mejora de la calidad del dato o en su producción: comedor, transporte, asistencia de docentes, etc
20	6. Demanda	Integrar en la planificación del sector algunas de las metas que plantea la sociedad civil, para aumentar la colaboración

15 - <http://opendatacharter.net/>

G.1.

RECOMENDACIONES SOBRE DATOS

	ÁMBITO	RECOMENDACIONES
PORTAL DE DATOS ABIERTOS		
21	5. Datos	Depurar duplicidades de datasets publicados por entidades territoriales en la categoría de infraestructuras educativas. Utilizar como referencia la información del DUE (Directorio de Establecimientos Educativos)
22	5. Datos	Vincular al portal de datos abiertos los servicios WMS que sirve la aplicación SI-GEO
23	5. Datos	Federar datasets que publican las iniciativas de Bogotá, Tuluá, ESRI o la red "Como vamos"
24	5. Datos	Revisar los sistemas de acceso a datos para mejorar la facilidad de acceso, evitando la necesidad de registro y permitiendo la descarga de fuentes de datos completas
IDENTIFICACIÓN DE NUEVOS DATASETS		
25	5. Datos	Incrementar la disponibilidad de datos económicos (presupuesto, gasto e inversión)
26	5. Datos	Incrementar la disponibilidad de datos de soporte institucional (becas, dotación para infraestructura de aprendizaje y enseñanza, etc.)
27	5. Datos	Incrementar la disponibilidad de datos de planta docente en educación preescolar, básica y media.
28	5. Datos	Disponer datos de los contratos de servicios que realiza MEN para el suministro de raciones alimenticias
29	5. Datos	Estudiar la publicación de la gestión financiera del PAE (Programa de alimentación escolar) a partir de los datos que procesa el sistema CHIP o SIIF.
30	5. Datos	Utilizar el análisis preliminar de disponibilidad de ficheros de datos para confeccionar un inventario de datos abiertos acorde a la guía de apertura de datos.
31	5. Datos	Determinar los orígenes de datos que tienen el mayor grado de actualización y precisión y utilizar éstas como fuente de datos principal.
32	5. Datos	Construir datasets a partir de los más de 600 archivos Excel identificados.
33	5. Datos	Disponer de datos en formatos descargables de los sitios web que están publicando datos de forma estructurada en formato HTML (PE: MIDE, datos de inspección y vigilancia, buscando carrera, etc.).
34	5. Datos	Incorporar como datos abiertos los servicios RSS.
35	5. Datos	Incorporar como datos abiertos los servicios WMS servidos por IGAC
36	5. Datos	Detectar de manera temprana las oportunidades de uso de los datos del futuro marco nacional de cualificaciones
37	5. Datos	Levantar un inventario de infraestructuras educativas que incluya información acerca de datos de servicio (programas, horarios...), con geoposición.
SISTEMAS DE INFORMACIÓN		
38	5. Datos	Estudiar la oportunidad de apertura de datos en los sistemas identificados como potenciales orígenes de datos
39	5. Datos	Estudiar la oportunidad de apertura de datos en los sistemas identificados como revisables
DATOS ENLAZADOS (Linked Data)		
40	5. Datos	Aprovechar las ontologías existentes (por ejemplo, la de Centros Educativos del Reino Unido), para impulsar una experiencia de publicación de datasets en formatos semánticos (RDF)

.ANEXOS

Anexo incluido en este documento:

ANEXO 1:

PERSONAS E INSTITUCIONES ENTREVISTADAS

Anexos presentados por separado a este documento:

ANEXO 2:

COMPARATIVO DATASETS INTERNACIONAL (HOJA DE CÁLCULO)

ANEXO 3:

IDENTIFICACIÓN DATASETS MINEDUCACIÓN (HOJA DE CÁLCULO)

ANEXO 4:

INVENTARIO SISTEMAS DE INFORMACIÓN MINEDUCACIÓN (HOJA DE CÁLCULO)

ANEXO 5:

INDICADORES INICIATIVA FUNDACIÓN CORONA (HOJA DE CÁLCULO)

ANEXO 6:

APPS EDUCACIÓN INTERNACIONAL (HOJA DE CÁLCULO)

ANEXO 1:

Personas e instituciones entrevistadas

#	SUBCATEGORÍA	ENTIDAD	CARGO
1	Entidades Públicas	Ministerio de Educación	Jefe de la Oficina de TI
2	Entidades Públicas	Ministerio de Educación	Líder GEL
3	Entidades Públicas	Ministerio de Educación	Coordinadora del grupo de aplicaciones
4	Entidades Públicas	Ministerio de Educación	Profesional Especializado
5	Entidades Públicas	Ministerio de Educación	Jefe Oficina de Innovación Educativa con Uso de Nuevas Tecnologías
6	Entidades Públicas	Ministerio de Educación	Subdirectora de Referentes y Evaluación de la Calidad Educativa del
7	Entidades Públicas	Ministerio de Educación	Director de la Calidad para la Educación Superior
8	Entidades Públicas	Ministerio de Educación	Subdirección de Desarrollo Sectorial
9	Entidades Públicas	Ministerio de Educación	Jefe Oficina Asesora de Planeación y Finanzas
10	Entidades Públicas	ICFES	Directora de Tecnología e Información del ICFES.
11	Entidades Públicas	ICFES	Subdirector de Información del ICFES.
12	Entidades Públicas	Ministerio de Educación	Coordinador del sistema de información
13	Sociedad civil	La Fundación Empresarios por la Educación (ExE)	Directora Ejecutiva
14	Sociedad civil	La Fundación Empresarios por la Educación (ExE)	Subdirectora de Programas y Proyectos
15	Entidades Públicas	Secretaría de Educación	Subsecretario de Calidad y Pertinencia
16	Entidades Públicas	Secretaría de Educación	Director de Ciencias, Tecnología y Medios Educativos
17	Entidades Públicas	Secretaría de Educación	Asesor del Secretario de Educación
18	Entidades Públicas	ICETEX	Director de Operaciones y Tecnología
19	Entidades Públicas	Ministerio de Educación	Jefe Oficina Asesora de Comunicaciones
20	Entidades Públicas	Ministerio de Educación	Unidad de Atención al Ciudadano
21	Entidades Públicas Adscrita a MinEducación	Colombia Aprende	Administrador Portal Colombia Aprende
22	Sociedad Civil	Fundación Corona	Subgerente Técnico
23	Sociedad Civil	Fundación Compartir	
24	Sociedad Civil	Fundación Compartir	Gerente de Educación
25	Sociedad civil	Fundación Dividendo por Colombia	Director de Programas



CATÁLOGO DE PRÁCTICAS INTERNACIONALES EN DATOS ABIERTOS DEL SECTOR EDUCACIÓN

BASE DE INSPIRACIÓN
PARA LA INICIATIVA
DE DATOS ABIERTOS
EN EDUCACIÓN EN
COLOMBIA



NOVIEMBRE DE 2015



MINTIC

vive digital
para la gente



GRUPO BANCO MUNDIAL



ÍNDICE

A. CATÁLOGO DE BUENAS PRÁCTICAS INTERNACIONALES **73**

A.1. Educación para todos y calidad educativa	75
A.2. Eficiencia en la gestión	79
A.3. Adecuación entre estudiantes y programas educativos	81
A.4. La vinculación entre los datos y la actividad empresarial y la adecuación de la enseñanza al empleo	83
A.5. Los datos, contextualizados, permiten avanzar hacia la web semántica o web de datos	87



A

CATÁLOGO DE BUENAS PRÁCTICAS INTERNACIONALES

EN este documento se recogen una serie de experiencias realizadas en base a datos abiertos en distintas iniciativas Open Data. Esta compilación de buenas prácticas demuestra que Educación es un sector en el que se está despegando y desarrollando una labor de reutilización de sumo interés social tan importante como ocurre en otros sectores como la movilidad, por ejemplo. Existen aplicaciones y servicios innovadores que abarcan desde la localización de equipamientos educativos, sus características y niveles de calidad, pasando por aplicaciones que dan soporte a la auditoría ciudadana, hasta llegar a una serie de implementaciones basadas en datos que se producen y abundan más en ámbitos académicos aprovechando

su mayor conocimiento técnico sobre analítica y combinación inteligente de los datos.

Como apoyo a este texto, se ha preparado un archivo Excel, que recoge las aplicaciones relacionadas con Educación desarrolladas en Reino Unido y USA, como iniciativas más emblemáticas a nivel mundial, así como las referenciadas en el Open Data Apps Directory¹ que gestiona el Oxford Internet Institute, que es un catálogo de aplicaciones que referencia más de 500 aplicaciones de todo tipo desarrolladas con datos abiertos a nivel mundial.

**>> Toda la información en
Hoja de Cálculo:
Apps_Educacion_Internacional.XLS**

1 - Open Data Apps: <http://opendata.oii.ox.ac.uk/>



ACCESO ≠
EDUCACIÓN
DE
CALIDAD

FOCO
ACCESO
EDUCACIÓN
CALIDAD

A.1

EDUCACIÓN PARA TODOS Y CALIDAD EDUCATIVA

LOS DATOS ABIERTOS APORTAN INFORMACIÓN FUNDAMENTAL PARA LOCALIZAR Y COMPARAR LA OFERTA ENTRE LAS INSTITUCIONES DE EDUCACIÓN

Es habitual utilizar datos de fuentes autorizadas de Gobierno para mostrar la disponibilidad y calidad de los centros educativos, detectar carencias de recursos educacionales a través de plataformas de mapas y a la vez ser la materia prima necesaria para la auditoría ciudadana y la colaboración para la mejora y el impulso a la innovación. Colombia ya está haciendo este ejercicio a través de aplicaciones

como “Buscando Colegio”, “Buscando Carrera” o SI-GEO.

Las herramientas que combinan orígenes de datos diversos, como por ejemplo: densidad de población, acceso al transporte, localización de centros de salud, redes de banda ancha, y otros factores sociodemográficos, son las utilidades sociales que tienen un mayor campo de acción y presentan una mayor variedad.

A CONTINUACIÓN SE DESCRIBEN ALGUNAS DE LAS MÁS RELEVANTES:

1. MEJORA TU ESCUELA (MÉXICO), es una plataforma para promover la participación ciudadana con el objetivo de mejorar la educación en México. La motivación del proyecto es el convencimiento de que la educación mejorará con el compromiso activo de todos los miembros de la comunidad educativa, en particular los padres de familia.

MejoraTuEscuela.org invita a la ciudadanía a buscar y conocer cómo está la escuela a la que acuden los niños, compararla con otras escuelas de una zona, calificarla y aportar opinión sobre las cosas que necesitan mejorar y las que ya se están haciendo bien. El proyecto proporciona herramientas para que los usuarios se vuelvan miembros activos y comprometidos para gestionar cambios positivos y mejoras en la comunidad educativa.

Promotores: Instituto Mexicano para la Competitividad A.C. (IMCO) con apoyo de la fundación Omidyar Network.

<http://www.mejoratuescuela.org/>

2. QEDU (BRASIL) herramienta cuyo desarrollo e implementación está financiado por la Fundacao Lemann, y que busca mejorar la calidad del aprendizaje en la educación brasilera. Tiene como objetivo estratégico estimular el debate público entre periodistas, maestros, familias y alumnos, ya que es la plataforma de datos educacionales más grande de Brasil. El Estado es el que alimenta los datos a este sistema a través del Ministerio de Educación; el sistema los procesa y publica en Internet en el portal www.qedu.org.br. Uno de los pilares del sistema es que presenta el grado de avance en educación, clasificado por temáticas y por estados. Esta clasificación permite conocer el contexto educativo a partir del nivel de aprendizaje en cada materia y por ciudad.

<http://www.qedu.org.br/>

Las dos aplicaciones anteriores se describen ampliamente en el estudio² “Datos abiertos en educación, primeros alcances y lecciones” (2015), elaborado por Yacine Khelladi para la Iniciativa Latinoamericana por los datos abiertos³.

2 - Estudio “Datos abiertos en educación, primeros alcances y lecciones”: <http://datosabiertos.org/wp-content/uploads/2015/10/6.Educacion-Khelladi.pdf>

3 - ILDA: Educación y datos abiertos: <http://datosabiertos.org/educacion-y-datos-abiertos/>

3. GREATSCHOOLS es el referente de la aplicación Mejora tu escuela en USA, con funcionalidades de relación social adicionales para reforzar la comunidad educativa

>> <http://www.greatschools.org/>

4. EDUCATION DASHBOARD (TANZANIA), es un portal público de datos abiertos a nivel nacional en forma de cuadro de mando, que permite acceso, visualización y monitoreo de los resultados de calidad educativa.

>> <http://educationdashboard.org/>

5. SHULE.INFO TANZANIA, es un proyecto Web que recopila los datos sobre los resultados de los exámenes de monitorio de la calidad educativa y las presenta en forma de visualizaciones.

Promotor: Twaweza [www.twaweza.org], organización de la sociedad civil con el soporte de diversos donantes internacionales.

>> <http://www.shule.info/>

6. Cabe mencionar, además, dos iniciativas interesantes llevadas a cabo en Punjab (India) y Burkina Faso que disponen de videos de presentación de cada una de sus experiencias:

Aplicación “School Education” desarrollada en el marco de la iniciativa Open Data de Punjab (India)

>> <https://vimeo.com/132905239>

Aplicación “Our Schools, our data” de la iniciativa Open Data de Burkina Faso.

>> <https://vimeo.com/97404977>

7. MAPKIBERA. EN KENIA, el Gobierno nacional y varias ONGs están usando una plataforma de mapas para determinar aquellas áreas donde los recursos educativos presentan fuertes carencias⁴. La iniciativa se ha ido extendiendo para dar cabida a otros temas además de la Educación como son la seguridad, agua o la salud.

>> <http://mapkibera.org/>

8. COLLEGECOST. El Departamento de Educación de USA, ha desarrollado un portal web que además de ofrecer información sobre los colegios incluye datos sobre el coste que tiene para un alumno el acceso a la oferta educativa:

>> <http://collegecost.ed.gov/>

El mismo Departamento de la administración USA publica el portal Web “Studentaid” que permite explorar distintas carreras, elegir y presentar la solicitudes de ingreso a instituciones educativas, los tipos de ayuda económica disponibles, provenientes del gobierno y de otras fuentes y cuales son los requisitos para solicitarlas, entre otras cuestiones de interés para los estudiantes:

>> <https://studentaid.ed.gov/sa/es>

Aprovechando el conocimiento técnico en general y específicamente sobre tratamiento de datos, en el ámbito universitario se está creando un catálogo de aplicaciones y soluciones sencillas que facilitan el acceso a información de otro modo compleja. Algunos ejemplos son:

9. LA UNIVERSIDAD PABLO DE OLAVIDE ESPAÑA, ha desarrollado la aplicación web análisis de alumnos de nuevo ingreso:

>> <https://datos.upo.gob.es/dashboard/>

Vinculado a la geo-localización, se están desarrollando aplicaciones y soluciones que permiten localizar y gestionar oferta y demanda de recursos y servicios en centros de educación superior:



4 - Mapkibera: <http://www.mapkibera.org/blog/2014/07/03/making-education-information-available-to-all-in-kibera/>

5 - Linked Open Data Universidad de Münster: <http://lodum.de/>

6 - Datos abiertos de la Universidad de Münster: <http://data.uni-muenster.de/>

10. CAMPUSPLAN es un proyecto es un proyecto de código abierto desde el Instituto de Geomántica de la Universidad de Münster (Alemania). La aplicación se basa en gran medida en los datos abiertos en el contexto de Linked Open Data de la Universidad⁵ publicados en su iniciativa⁶ de datos abiertos.

>> <http://app.uni-muenster.de/>

11. En la misma línea, la **UNIVERSIDAD DE OXFORD**, ofrece los datos de todos los edificios del Campus por medio de una visualización interactiva geo-localizada.

>> <https://data.ox.ac.uk/explore/science-area/>

12. OPEN UNIVERSITY DEL REINO UNIDO, publica un banco de aplicaciones que usan datos abiertos del portal de la propia Universidad, que van desde la catalogación de recursos educativos, mapas de edificios o una aplicación que permite geo localización entre personas dentro del Campus.

>> <http://data.open.ac.uk/site/applications.html>

13. LA UNIVERSIDAD DE SOUTHAMPTON propone una aplicación que permite buscar y localizar salas de trabajo libres dentro del Campus de la Universidad. Room-finder, está diseñada para los miembros de la Universidad que necesitan, de forma inmediata, encontrar espacios de trabajo no reservados. Utiliza datos abiertos de los edificios, las características de las salas y los datos de gestión de reservas permitiendo al usuario buscar en estos espacios en base a su disponibilidad e idoneidad.

>> <http://tools.southampton.ac.uk/room-finder/>

Por último y relacionado con este aspecto, no se puede dejar de mencionar la gran oportunidad que incorporan los proyectos de ciudades inteligentes que se desarrollan en la actualidad y que van a permitir combinaciones muy ricas de distintos orígenes de datos.





A.2

EFICIENCIA EN LA GESTIÓN

OPEN DATA ES UNA OPORTUNIDAD PARA MEJORAR LA RESPONSABILIDAD GERENCIAL Y LA DETECCIÓN DE INEFICIENCIAS EN LA GESTIÓN DE LOS RECURSOS POR PARTE DE LOS CENTROS EDUCATIVOS.

Los medios, en los procesos de investigación periodística, utilizan habitualmente información gubernamental dispersa y de otras fuentes para denunciar casos de ineficiencias económicas financieras en las instituciones del estado. En muchos casos, la información con la que se trabaja es difícil de conseguir o contrastar, generando debates sociales que adolecen de falta de objetividad, porque se especula con la información, dando lugar a la movilización ciudadana para realizar auditoría social de la política pública de verificación de información básica. Sin menoscabar el valor de este tipo de iniciativas (tanto las realizadas por los medios de comunicación como por la sociedad en su conjunto), la apertura de datos permite un debate objetivo y creativo basado en hechos, lo que aporta un valor agregado al desarrollo social y económico del país: no debatimos sobre cuánto se asigna a una determinada institución educativa, sino que el debate se centrará en la eficiencia y eficacia del gasto. En ello juega un papel fundamental tanto el periodismo de datos como la ciudadanía inteligente y bien informada.

7 - ANSA-EAP: <http://www.ansa-eap.net/>

1. El sitio web **CHECKMYSCHOOL DE FILIPINAS** es una aplicación de auditoría ciudadana para revelar y detener las indicios de corrupción que pueden tener las escuelas al tratar de conseguir los fondos a los que tienen derecho. El departamento de Educación de ese país ha luchado con las ineficiencias, las prácticas corruptas y la educación de mala calidad. Desarrollado por la organización civil Affiliated Networks of Social Accountability (ANSA⁷), CheckMySchool consolida todos los datos disponibles del Gobierno sobre más de 40 mil escuelas públicas en las Filipinas. La web proporciona una plataforma interactiva para la retroalimentación ciudadana, con tres objetivos principales: la provisión de información, la validación de datos del gobierno y el de participación comunitaria.

2. LA UNIVERSIDAD DE ALICANTE ESPAÑA, a partir de su iniciativa de datos y al amparo de la Ley de Transparencia (19/2013, de la legislación española), realiza un interesante ejercicio de vinculación entre publicación activa de información sujeta a lo dispuesto en la legislación, y datos abiertos disponibles para su descarga y reutilización:

>> <http://transparencia.ua.es/es/>



A.3

ADECUACIÓN ENTRE ESTUDIANTES Y PROGRAMAS EDUCATIVOS

OPEN DATA ES UNA HERRAMIENTA EXTRAORDINARIA PARA COMPARTIR Y ENRIQUECER PROGRAMAS FORMATIVOS ADAPTADOS A LOS PERFILES DE LOS ESTUDIANTES Y ES ADEMÁS UNA VALIOSA HERRAMIENTA PARA EL APRENDIZAJE Y DESARROLLO DE HABILIDADES EN CIENCIA DE DATOS Y PROGRAMACIÓN

El acceso abierto, conocido con las siglas OA (Open Access), promueve que el conocimiento sufragado públicamente sea de dominio público y postula un acceso permanente, gratuito y libre de restricciones a los contenidos científicos y académicos. Este movimiento está avalado por distintas declaraciones internacionales.

LOS DATOS A LOS QUE SE REFIEREN ESTAS INICIATIVAS SON TÍPICAMENTE DE TRES CLASES:

- Material docente: OpenCourseWare (OCW⁸) y Open Educational Resources (OER⁹); y cursos abiertos Massive Open Online Course (MOOC)
- Las publicaciones y otros documentos que recogen los resultados de la investigación o informan sobre ella, recogidos normalmente en repositorios institucionales.
- Datos experimentales o técnicos sobre los que se realiza una investigación, para que pueda ser replicada por otros, o para que otros puedan usarlos para sus propias investigaciones. Existe un importante movimiento internacional en favor de la compartición eficaz de datos, vertebrado en la Research Data Alliance (RDA¹⁰).

Colombia tiene una oportunidad importante de sumarse a este movimiento poniendo a disposición en formatos estándar y abiertos su catálogo de recursos educativos que contiene más de 60 mil recursos.

1. LA OPEN KNOWLEDGE FOUNDATION¹¹, promueve el grupo de trabajo Open Education que va más allá de ser un repositorio y punto de referencia en OER y trata aspectos como la apertura de los datos educativos relevantes y el cambio institucional y cultural necesario.

Finalmente, para completar este apartado es necesario introducir la idea de que datos abiertos es un habilitador muy importante para las enseñanzas técnicas y el desarrollo de habilidades analíticas, infográficas y programación. Profesiones vinculadas a “big y small data”, analítica y minería de datos, gestores de recursos “cloud”, desarrolladores de aplicaciones móviles, o profesionales del arte digital, requieren competencias relacionadas con datos. Actualmente, existe una gran diversidad de cursos (muchos MOOC) relacionados con datos, pero además se pueden destacar estas iniciativas:

2. ESCUELA DE DATOS es una iniciativa promovida desde la Open Knowledge Foundation, que tiene como objetivo empoderar a organizaciones civiles, periodistas y ciudadanos para que sean capaces de usar los datos de manera efectiva y eficiente. La misión de Escuela de Datos es enseñarte a usar los datos como fuente de información para contar una historia.

>> <http://es.schoolofdata.org/>

3. TUALABS, es una empresa especializada en la enseñanza basada en datos abiertos. Su misión es hacer de los datos abiertos una materia prima útil para la enseñanza y el aprendizaje, ofreciendo a profesores, estudiantes y padres de familia una plataforma para analizar, visualizar e interpretar datos reales en torno a temas relevantes para sus comunidades e intereses. Esta entidad ha ayudado a promover habilidades para trabajar con datos en USA, India, y otros lugares, además de una colaboración con el Banco Mundial en Sudán.

>> <https://tvalabs.com/>

8 - OpenCourseWare: <https://es.wikipedia.org/wiki/OpenCourseWare>

9 - OER: https://es.wikipedia.org/wiki/Recursos_educativos_abiertos

10 - RDA: <https://rd-alliance.org/>

11 - Open Knowledge Foundation: <https://okfn.org/>



A.4

LA VINCULACIÓN ENTRE LOS DATOS Y LA ACTIVIDAD EMPRESARIAL Y LA ADECUACIÓN DE LA ENSEÑANZA AL EMPLEO

OPEN DATA PROMUEVE LA ACTIVIDAD ECONÓMICA Y EL DESARROLLO EMPRESARIAL EN BASE A DATOS ABIERTOS, IMPLEMENTANDO APLICACIONES RICAS COMBINACIÓN DE ORÍGENES DE DATOS DIVERSOS, CONTRIBUYENDO A LA MEJORA DE LA CALIDAD DE VIDA Y LA CREACIÓN DE EMPLEO CUALIFICADO.

1. “Opendata500” es un proyecto coordinado por GovLab¹² cuya misión es el estudio del uso e impacto de los datos abiertos. Se articula en forma de red internacional de promotores de iniciativas de datos abiertos que analizan el comportamiento de sus respectivos ecosistemas y permiten una comparación homogénea sobre el uso de los datos abiertos entre países y así mejor entender el impacto económico y social de las diversas iniciativas de datos abiertos. Los objetivos del estudio son:

Proveer una base para evaluar el valor económico y social de los datos gubernamentales abiertos.

Estimular el desarrollo de nuevas compañías que trabajan con Datos Abiertos.

Facilitar diálogo entre el gobierno y el sector privado sobre cómo mejorar la utilidad de los datos gubernamentales abiertos.

Colombia tiene una oportunidad importante de sumarse a este movimiento poniendo a disposición en formatos estándar y abiertos su catálogo de recursos educativos que contiene más de 60 mil recursos.

En estos estudios se presentan resultados que responden a la pregunta ¿Qué tipo de compañías utilizan datos provenientes de cuáles entidades?

En la iniciativa participan actualmente Australia, México, Estados Unidos, Italia y Korea.

>> <http://www.opendata500.com/>

Finalmente, para completar este apartado es necesario introducir la idea de que datos abiertos es un habilitador muy importante para las enseñanzas técnicas y el desarrollo de habilidades analíticas, infográficas y programación. Profesiones vinculadas a “big y small data”, analítica y minería de datos, gestores de recursos “cloud”, desarrolladores de aplicaciones móviles, o profesionales del arte digital, requieren competencias relacionadas con datos. Actualmente, existe una gran diversidad de cursos (muchos MOOC) relacionados con datos, pero además se pueden destacar estas iniciativas:

2. A partir de “Opendata500”, el estudio correspondiente a México se materializa en la iniciativa “Datos Abiertos 100 México” donde se presentan los resultados del primer estudio de las compañías y ONGs mexicanas que utilizan datos gubernamentales abiertos para crecer sus negocios, desarrollar nuevos productos y servicios, o impulsar el valor social.

Como resultado del estudio, y en el contexto del sector educativo los datos procedentes de la Secretaría de la Educación Pública son consumidos por compañías de: investigación y consultoría (7), tecnología/datos (5), servicios empresariales (2), servicios legales (1), vivienda/bienes raíces (1), salud (1) geoespacial/mapas (1).

Promotor: Equipo de la Estrategia Nacional Digital de la Oficina de la Presidencia de la República.

>> <http://www.opendata500.com/mx/>

12 - Thegovlab: <http://www.thegovlab.org/>

3. ODINE¹³, (Open Data Incubator Europe) es un proyecto financiado por programa de investigación y desarrollo Horizon 2020 de la Unión Europea, cuya misión es promover la generación de actividad empresarial sostenible a partir de datos abiertos. El proyecto, a través de diversas convocatorias, proporciona financiación y soporte a propuestas surgidas de pequeñas y medianas empresas y desde el emprendimiento.

En la última convocatoria, una de las siete startups beneficiadas con el programa de financiamiento es la que desarrolló la aplicación Sickly.

4. SICKLY reúne datos abiertos sobre la propagación de enfermedades infecciosas entre los niños a partir de los datos que proporcionan los padres que utilizan una aplicación gratuita y pueden reportar de forma segura a la escuela la enfermedad que sus hijos padecen en un determinado momento. El objetivo de Sickly es rastrear enfermedades infecciosas utilizando datos anónimos y servir de apoyo a los centros de salud pública en la lucha contra las enfermedades infecciosas.

>> <https://www.sickly.org/>

5. VACANCY TREEMAP. La Universidad de Oxford ha desarrollado una interesante aplicación que vincula la oferta educativa con la oferta de empleo por medio de una visualización que relaciona cada departamento de la universidad con las vacantes de empleo (público o privado) relacionada con esa especialidad.

>> <https://data.ox.ac.uk/explore/vacancy-treemap/>

6. En esta categoría de experiencias es imprescindible mencionar, las iniciativas surgidas desde el ámbito público para el estímulo de la actividad reutilizadora y que además implican al entorno académica. Un ejemplo de ello es la Iniciativa Open Data Terrassa (España), de ámbito municipal. En el video enlazado a continuación se destaca la importancia de la vinculación de Open Data con el proceso educativo (Centros escolares y universidad) donde se conjuga la oferta de datos (Ayuntamiento), el desarrollo de aplicaciones (centros educativos), la orientación y dirección técnica (universidad), y la llegada al mercado con el soporte de las herramientas de desarrollo económico de la administración municipal.

>> <https://vimeo.com/66721518> (minuto 4':47")

Cabe mencionar en este apartado, uno de los mecanismos tradicionales en las estrategias de fomento de la reutilización, son los concursos o maratones de datos y datathons que juegan un papel fundamental en todas las incitativas de datos abiertos estimulando el consumo, promoviendo la innovación y facilitando la interacción entre los miembros de la comunidad Open Data. Las versiones evolucionadas de estas maratones de desarrollo tienen que ver con la co-creación y son procesos más continuados en el tiempo, más participativos y colaborativos, más guiados, generalmente orientados hacia una temática concreta; contando con la participación de expertos en esa área; con nuevos datos disponibles para la ocasión; y además con algún tipo de incentivo que va más allá del simple premio buscando la generación de modelos de negocio realistas a medio plazo. Existen ejemplos con estas características:

7. LINKED UP CHALLENGE, centrado en la temática de los datos de propósito educativo y dividido en tres fases consecutivas.

>> <http://linkedup-challenge.org/>

8. LA SERIE OPEN DATA CHALLENGES, organizada por Nesta y el Open Data Institute, y que hasta el momento ha cubierto ya temáticas como Seguridad y Justicia; Energía y Medio Ambiente; Educación; y Alimentación, y con un especial foco en la generación de modelos de negocio sostenibles para las soluciones propuestas.

>> <http://www.nesta.org.uk/project/the-open-data-challenge-series>

9. LOS EUROPEANA CREATIVE CHALLENGES, donde se exploran las posibles aplicaciones de los datos culturales en campos como la Educación o el Turismo.

>> <http://www.europeanacreative.eu/web/europeana-creative/challenges>

13 - Proyecto ODINE: <https://opendataincubator.eu/eu/overview/>





A.5

LOS DATOS, CONTEXTUALIZADOS, PERMITEN AVANZAR HACIA LA WEB SEMÁNTICA O WEB DE DATOS

OPEN DATA ES UNA DE LAS BASES QUE CONSTITUYEN EL FUTURO DE LA WEB. UN CAMINO QUE HAY QUE RECORRER DESDE LA INICIATIVA PÚBLICA CON LA IMPLICACIÓN ACTIVA DE LA COMUNIDAD ACADÉMICA Y EL SECTOR EMPRESARIAL.

La Web Semántica es una extensión de la web actual en la cual se le da a la información un significado bien definido, lo que permite a las computadoras y a las personas trabajar en cooperación¹⁴. Para entender el concepto es fundamental entender el funcionamiento de la web actual que actúa como una biblioteca de documentos (páginas web multimedia) interconectadas por enlaces, o lo que es lo mismo, un espacio en el que las computadoras son responsables de la representación (la parte más sencilla) y las personas se encargan de la interpretación (lo más complicado). ¿Se puede invertir la complejidad o al menos que las computadoras hagan tareas más difíciles? Es posible, y para ello es necesario enriquecer la información disponible en las páginas web con semántica (significado) que las máquinas puedan entender.

Los elementos clave para hacerlo son las ontologías y las anotaciones. Las ontologías, que normalmente se representan con lenguajes como RDF Schema¹⁵ u OWL¹⁶, describen formalmente conceptualizaciones compartidas de un dominio (por ejemplo, personas, reuniones, centros escolares, etc.) Las anotaciones, que se describen normalmente en RDF, permiten describir instancias de dichas ontologías y asociarlas a recursos Web (por ejemplo, diciendo que la información contenida en una página web se refiere a un centro escolar). Esto permite proporcionar acceso inteligente a información heterogénea y distribuida, permitiendo que las aplicaciones software puedan intermediar entre las necesidades de los usuarios y las

fuentes de información disponibles. Una forma gráfica y didáctica de entender el concepto, se encuentra en este enlace¹⁷.

Por otro lado, el concepto de Linked Data se refiere al mecanismo por el cual se disponen públicamente datos en RDF mediante la utilización del protocolo http (a través de un identificador del dataset único llamado URI), es decir, son una manera normalizada de representar datos sobre una gran variedad de temas. La publicación de datos vinculados o Linked Data facilita a los programadores la conexión de información de diferentes fuentes, lo que lleva a nuevas e innovadoras aplicaciones. Por ejemplo, una aplicación puede combinar datos sobre sectores empresariales predominantes, emisiones al medio ambiente o centros educativos. Este concepto ha ganado fuerza después de la publicación de recursos como DBpedia¹⁸, Geonames¹⁹, etc., y por el impulso de varios gobiernos con su decisión de hacer públicos sus datos abiertos utilizando esta tecnología, como es el caso de data.gov (USA), data.gov.uk (UK) o datos.gob.es en España.

No es objeto de este documento avanzar en el alcance de la web semántica, tema sobre el que existe abundante bibliografía, en cambio si es objetivo identificar recursos útiles para implementar experiencias en base a datos enriquecidos semánticamente en el ámbito educativo. La mayoría de los recursos que se exponen a continuación tienen que ver con la localización de infraestructuras educativas o la compilación de objetos de aprendizaje.

14 - Definición del artículo publicado por Tim Berners Lee, James Hendler y Ora Lassila en 2001 en la revista Scientific American.

15 - RDF Schema es una extensión de RDF que permite expresar enunciados sobre recursos usando propiedades y valores.

16 - OWL es el lenguaje estándar de la web semántica para expresar y codificar ontologías.

17 - <http://www.ukma.edu.ua/~gor/Agents/Refs/PDFTexts/SemaviewThe%20Semantic%20Web.pdf>

18 - DBpedia: <http://wiki.dbpedia.org/about>

19 - Geonames: <http://www.geonames.org/>

LA INICIATIVA DE DATOS ABIERTOS CON MAYOR EXPERIENCIA Y RECORRIDO EN LINKED DATA Y SU APLICACIÓN EN EL SECTOR DE EDUCACIÓN ES LA DEL REINO UNIDO QUE HA IMPLEMENTADO LAS SIGUIENTES REFERENCIAS

ONTOLOGÍAS DE UK PARA EL SECTOR EDUCATIVO

<http://source.data.gov.uk/def/education/>

LINKED DATA API PARA BÚSQUEDA DE ESCUELAS

<http://education.data.gov.uk/def/school/>

ONTOLOGÍAS DE UK PARA OTROS SECTORES

<http://source.data.gov.uk/def/>

USANDO DATOS DE EDUCACIÓN DE UK

<http://openuplabs.tso.co.uk/datasets/gov-education>

1. EL CATÁLOGO DE DATOS Y APLICACIONES DE LA UNIVERSIDAD DE SOUTHAMPTON es uno de los casos de uso más sobresaliente en el campo de Linked Data aplicado en el ámbito de la educación superior. En su amplio catálogo de datos se pueden encontrar cursos académicos, edificios y localizaciones, menús diarios, repositorios documentales, etc. disponibles principalmente en formato RDF y otros formatos procesables por máquinas. Las aplicaciones que consumen los datos son directorios, mapas, localización de buses, equipamientos o catering disponibles en forma de visualizaciones web, widgets o aplicaciones móviles.

>> <http://data.southampton.ac.uk/>

2. LINKED EDUCATION, es un proyecto que compila un repositorio de datasets relevantes para el desarrollo de aplicaciones educativas. Su principal característica es que responde al estándar de la Web de datos y se construye sobre la base de las aportaciones de la comunidad LinkedUp²⁰, acción financiada por la Comisión Europea para impulsar la explotación y adopción de políticas Open Data en las organizaciones e instituciones educativas en Europa.

>> <http://data.linkededucation.org/>

3. DIDACTALIA, es una comunidad educativa global para profesores, padres y estudiantes desde la Educación Infantil hasta el Bachillerato que incluye una colección con más de 50.000 recursos educativos abiertos. En ella, los usuarios pueden crear, compartir y descubrir contenidos educativos y promover clases y comunidades de aprendizaje en las que trabajar lecciones en un espacio social con búsquedas facetadas y contextos enriquecidos.

En cuanto a identidad digital, Didactalia.net propone un Entorno Personal de Aprendizaje (PLE), que incluye herramientas semánticas y sociales para gestionar la reputación y un espacio en el que el usuario recoge, expone y comparte sus contenidos de aprendizaje. El proyecto está liderado por la empresa GNOSS.

>> <http://datahub.io/dataset/didactalia>

4. REPOSITORIO DE VOCABULARIOS LINKED OPEN VOCABULARIES DE LA OPEN KNOWLEDGE FOUNDATION, que recoge información y acceso a 527 vocabularios entre los que hay una categoría vinculada al mundo académico.

>> <http://lov.okfn.org/dataset/lov>

5. DATAHUB, compila recursos de acceso libre a plataformas de datos desarrolladas sobre CKAN y permite la búsqueda y gestión de datasets. La iniciativa está liderada por la Open Knowledge Foundation. Entre otras categorías mantiene una sobre Educación.

>> http://datahub.io/organization/ce086ef6-9e56-4af1-a63c-fd768fa2dfff?tags=education&_tags_limit=0

6. LINKED UNIVERSITIES, es una alianza de Centros de Educación Superior europeas implicadas en la publicación de datos abiertos como Linked Data. En su portal Web publica directorio de catálogos de datos de los Centros de Educación Superior miembro, vocabularios y ontologías, endpoints de consulta, diversas herramientas y publicaciones.

>> <http://linkeduniversities.org/index.html>

20 - Linking web data for Education: <http://linkedup-project.eu/>



**MÁS DETALLE DE OTRAS APLICACIONES, EN EL ANEXO:
Apps_Educacion_Internacional.XLS**