

Problemáticas en la difusión de datos educativos. El tratamiento de su confidencialidad en Argentina.

De Grande, Pablo.

Cita:

De Grande, Pablo (2025). *Problemáticas en la difusión de datos educativos. El tratamiento de su confidencialidad en Argentina*. e-ISSN: 2362-3349 - Revista de la Escuela de Ciencias de la Educación, 20 (1), 100-117.

Dirección estable: <https://www.aacademica.org/pablo.de.grande/81>

ARK: <https://n2t.net/ark:/13683/pcWP/7xV>



Esta obra está bajo una licencia de Creative Commons.
Para ver una copia de esta licencia, visite
<https://creativecommons.org/licenses/by-nc-nd/4.0/deed.es>.

Acta Académica es un proyecto académico sin fines de lucro enmarcado en la iniciativa de acceso abierto. Acta Académica fue creado para facilitar a investigadores de todo el mundo el compartir su producción académica. Para crear un perfil gratuitamente o acceder a otros trabajos visite: <https://www.aacademica.org>.

Problemáticas en la difusión de datos educativos. El tratamiento de su confidencialidad en Argentina¹

Problems in the dissemination of educational data and the treatment of its confidentiality in Argentina

Pablo de Grande

pablodg@gmail.com

<http://orcid.org/0000-0003-1245-1573>

Universidad del Centro de la Provincia de Buenos Aires, Instituto de Estudios Histórico Sociales (IEHS-IGEHCS); CONICET, Argentina

Resumen

El artículo propone evaluar la modalidad de publicación de datos de la Secretaría de Evaluación e Información Educativa del Ministerio de Educación de la Nación de Argentina. Se examina la accesibilidad y la confiabilidad alcanzados por la misma, a partir de la revisión de resoluciones de reclamos de acceso a la información pública y de un análisis del esquema de protección que el Ministerio aplica sobre la información publicada. Como principales resultados se destaca un escenario de sobreprotección de los datos –en términos de excluir del acceso a información pública relevante a expertos y al público en general– combinada con una llamativa debilidad del esquema de protección, que no sigue ningún estándar probado y que expone la información considerada confidencial de manera irrestricta. Esta investigación puede facilitar ejercicios comparativos entre países de la región con el propósito de enriquecerse de las experiencias de cada uno, aprender de sus errores y sugerir caminos de mejora para la puesta a disposición de información educativa.

Palabras clave: Información educativa, gobierno abierto, confidencialidad, confiabilidad, accesibilidad.

Abstract

The article evaluates how data is published by the Secretaría de Evaluación e Información Educativa of the Ministerio de Educación de la Nación in Argentina. The accessibility and reliability are examined, based on the review of resolutions of claims for access to public information and an analysis of the protection scheme that the Ministerio applies on the published information. The main results highlight a scenario of overprotection of data –in terms of excluding experts and the general public from access to relevant public information– combined with a striking weakness of the protection scheme, which does not follow any proven standard and exposes information considered confidential in an unrestricted manner. This research can facilitate comparative exercises between countries in the region with the purpose of enriching each other's experiences, learning from their mistakes and suggesting paths for improvement in making educational information available.

Keywords: Educational information, open government, confidentiality, reliability, accessibility

Introducción

La educación pública en Argentina vive una crisis crónica: mientras que los salarios docentes son recurrentemente noticia por su insuficiencia (El cronista, 2023; Guadagni, 2018), la opinión negativa sobre la situación de la educación es mayoritaria entre los residentes del país (IGEDECO, 2023) y las evaluaciones sobre el rendimiento estudiantil son desalentadoras (Chávez et al., 2022). Estos fenómenos ocurren a la par de infructuosos esfuerzos para mejorar la terminalidad en el nivel secundario (De Grande, 2019) y de un desplazamiento del foco en este nivel de lo ‘pedagógico’ a lo ‘socioeducativo’ (Tiramonti, 2019).

Este estado de cosas amplía las desigualdades existentes entre quienes acceden al sistema de educación estatal y quienes acceden a establecimientos de gestión privada. Según los informes de resultados de pruebas estandarizadas del año 2022, en el nivel primario 76,1% de los niños en establecimientos de educación privada alcanzaron niveles satisfactorios y avanzados en matemática, mientras que solamente el 51,2% lograron esos niveles en establecimientos estatales (Dirección de Información y Estadística Educativa (DIEE), 2023b, p. 106)².

En este contexto, la necesidad de diagnósticos que ayuden a la comprensión de estas problemáticas es por demás evidente. Su persistencia en el tiempo, unida al carácter estratégico del desarrollo de la educación para la recuperación social y económica del país, conducen a poner el foco en la información con que cuentan quienes gestionan, investigan y se interesan en la educación y sus posibilidades de mejora.

Este artículo indaga en el problema del acceso a la información educativa en Argentina, buscando poner en cuestión tanto el aspecto restrictivo en vigencia del mismo (la negativa del Ministerio de Educación a facilitar información educativa) como la modalidad en que dichas restricciones se refuerzan.

Para esto, se hará en primer lugar una síntesis de la situación de la información educativa en la Argentina, considerando su marco normativo y las limitaciones existentes en el acceso. En segundo lugar será analizada, dentro de ese contexto,

la modalidad utilizada por la Secretaría de Evaluación e Información Educativa del Gobierno Nacional para el resguardo de la privacidad en la publicación de datos, mostrando las debilidades técnicas de esta estrategia. Finalmente, se retoman las preocupaciones iniciales para hacer una síntesis de los problemas encontrados en la accesibilidad (poco acceso) y en la confiabilidad (nula protección) de la información educativa.

Desarrollo

Marco legal para la publicación de datos educativos en Argentina

La publicación de datos educativos en establecimientos estatales debe atender a la necesidad de transparencia y monitoreo de la gestión pública, resguardando en el proceso la privacidad de sus actores (en particular de los estudiantes). En Argentina, el marco legal que la regula es complejo y parcialmente contradictorio.

Por un lado, la Ley 27.275 de Acceso a la Información Pública, sancionada en 2016, establece el principio de “transparencia activa”, que obliga a los organismos estatales a facilitar el acceso a la información que producen y de que disponen. Por otro lado, la Ley 17.622 de Estadísticas y Censos, del año 1968, establece restricciones en la publicación de datos individuales recolectados con fines estadísticos, a la vez que la Ley 26.206 de Educación Nacional, del año 2006, establece criterios particulares para la publicación de resultados de evaluaciones.

En este contexto, Argentina enfrenta desafíos significativos en la gestión y publicación de su información educativa. A diferencia de otros países de la región, carece de un sistema de información que permita realizar un seguimiento de las trayectorias individuales de sus estudiantes (Bonelli, 2021). Adicionalmente, la información disponible sobre las escuelas es limitada, ya que se publica sin permitir identificar a los establecimientos, lo que dificulta el análisis del sistema educativo en su conjunto (Suasnábar y Valencia, 2022).

Dos elementos son llamativos en estas tensiones. Por un lado, la Secretaría de Evaluación e Información Educativa del Ministerio de Educación aplica criterios de protección de privacidad y confidencialidad que, si bien podrían ser pertinentes para proteger a personas individuales, al

aplicarse sobre establecimientos —y en particular sobre aquellos de gestión pública— da por resultado volver opaco el funcionamiento del Estado en la provisión de servicios educativos. Sorprende en este aspecto que, al considerar el grado de desagregación con que facilita la información, la Secretaría de Evaluación e Información Educativa no distinga entre organismos de gestión estatal (cuyo funcionamiento debe someterse a escrutinio abierto, por gestionarse plenamente con fondos públicos) respecto de las instituciones de gestión privada (a las que pueden caberle derechos de confidencialidad específicos de su actividad).

Por otro lado, si bien la publicación de datos con resguardo de privacidad en el contexto internacional se apoya en discusiones y saberes de varias décadas (Samarati y Sweeney, 1998; Esquivel-Quirós et al., 2019), se implementa en la información educativa de Argentina de manera sorprendentemente rudimentaria.

Sobre este último punto, cabe señalar que los estándares internacionales en esta área enfatizan la importancia de encontrar un equilibrio entre la preservación de la privacidad y el aprovechamiento del conocimiento que puede derivarse de los datos. Esto implica no solo publicar resultados tabulados, sino también dar acceso a los registros subyacentes, aplicando técnicas de anonimización (Fung et al., 2010).

Sin embargo, las prácticas actuales de la Secretaría de Evaluación e Información Educativa permanecen alejadas de estos estándares. La estrategia basada en “identificadores de fantasía” para proteger la identidad de las escuelas, como detalla más adelante este artículo, falla en garantizar su anonimización, a la vez que impide un razonable equilibrio entre accesibilidad y protección de la privacidad.

Planteamiento del problema

La publicación de bases de datos educativas de alcance nacional en el país se realiza principalmente por medio de cuatro instrumentos de publicación periódica: el padrón de establecimientos educativos (que permite listar los establecimientos existentes por nombre y algunos atributos generales), el mapa educativo (que permite georreferenciarlos, sin actualización desde 2019), el relevamiento anual de establecimientos (que capta información detallada de

alumnos, docentes, secciones, instalaciones, entre otros), y los resultados de las pruebas estandarizadas Aprender (antes ONE, por Operativo Nacional de Evaluación).

El mapa educativo y el padrón de escuelas incluyen el nombre de cada establecimiento y su código único de identificación (cueanexo), por lo que es sencillo vincular elementos de un listado con el otro. Sin embargo, las evaluaciones educativas y los relevamientos anuales (allí donde se encuentra casi la totalidad de la información publicada) son anonimizadas para su publicación. En consecuencia, si bien se publican sus microdatos, los mismos no pueden vincularse ni con la escuela que describen, ni tampoco con niveles administrativos básicos como el distritito escolar, el municipio o el departamento de pertenencia. En estos dos conjuntos de datos en lugar de incluirse para cada fila su identificador de establecimiento (cueanexo), se utiliza un identificador “enmascarado” llamado “Id de fantasía”.

Tal como señala la bibliografía especializada, la discontinuidad existente entre las cuatro fuentes de información es problemática (Suasnábar y Valencia, 2022). Los establecimientos actuales pueden listarse, pero no mapearse (el mapa está desactualizado). Los relevamientos y las evaluaciones no pueden ni agruparse por niveles administrativos subprovinciales (distrito escolar, municipio, etc.) ni vincularse con datos censales a través de su ubicación espacial para producir diagnósticos territorializados o relacionando las condiciones socioeconómicas y demográficas de contexto de cada institución educativa.

En junio de 2023, 64 organizaciones junto a 53 expertos en investigación educativa presentaron una solicitud al Ministerio de Educación para que hiciera pública información educativa básica del país (cantidad de alumnos, docentes, etc.) a nivel departamental/municipal, apoyándose en las normativas de acceso a la información pública vigentes y en la imperiosa necesidad de progresar en los análisis de nuestra realidad educativa. El Ministerio se negó a facilitar la información, citando al pasaje de la ley que regula las estadísticas oficiales donde se afirma que “los datos deberán ser suministrados y publicados, exclusivamente en compilaciones de conjunto, de modo que no pueda ser violado el secreto comercial o patrimonial, ni individualizarse las personas o entidades a quienes se refieran” (AAIP, 2023, p. 2).

Al restringir el acceso a la información de escuelas de gestión pública, el Ministerio considera que debe aplicar la misma protección para estos establecimientos estatales que la prevista en la ley para proteger a persona particulares de cara a un relevamiento estadístico (como un censo de población).

Esta disposición pone en evidencia la incapacidad del Ministerio (continuada a lo largo de gestiones de diferente signo político) de reconocer el enorme problema que su modalidad de publicación de datos representa para la comunidad de usuarios interesados en analizar temáticas educativas. De otro modo, pondría a disposición canales para que la información educativa sea accedida. Entre ellos, podrían existir —además de una revisión integral de la interpretación del cuerpo legal que aplica— acuerdos de confidencialidad previos a la entrega de las bases de datos con identificadores reales (que comprometan a los receptores a no publicar información por debajo de cierto nivel de desagregación); apertura a realizar procesamientos por solicitud de usuarios; circuitos para acreditar el acceso a herramientas de consulta avanzadas a especialistas; publicación de datos en formato Redatam (el formato utilizado por el Instituto de Estadísticas y Censos); entre otros.

La imposibilidad de consistir la información educativa con sus localizaciones geográficas y político-administrativas, y la recurrente negativa de la Secretaría de Evaluación e Información Educativa del Ministerio de Educación a responder a solicitudes de pedidos de información en aspectos tan elementales como la cantidad de alumnos, docentes o aulas, abre interrogantes cuya discusión parece necesaria.

Por un lado, ¿es adecuado el volumen de información que se considera confidencial? En el caso del relevamiento anual de escuelas, se trata de 763 atributos por establecimiento considerados, en su totalidad, confidenciales. En el caso de las pruebas estandarizadas, abarca desde el rendimiento en matemáticas y lengua, hasta las preguntas de opinión hechas a los estudiantes en el cuestionario de cántula.

Esta primera pregunta será abordada en este artículo por medio del análisis de resoluciones de la Agencia de Acceso a la Información Pública, las cuales muestran que el Ministerio se encuentra en contradicción con la normativa y la jurisprudencia

vigentes, las cuales propician un equilibrio entre la apertura de datos y la defensa de la confidencialidad bien diferente al sostenido por el Ministerio.

En segundo lugar, para dar integralidad al análisis emprendido, cabe preguntarse —en caso de que la información publicada fuera en efecto confidencial— si es adecuado el modo en que se protege. La estrategia sostenida para los datos del período 2012-2022 es distribuir los microdatos acompañados de identificadores ficticios para cada establecimiento. Sobre esta segunda pregunta, el artículo realiza un análisis al tipo de protección que se ofrece, arribando a la conclusión de que no se utiliza ningún estándar criptográfico probado, ni se resguarda la información de manera efectiva, por lo que parece necesaria una intervención de importancia si existiera allí información confidencial a proteger.

El acceso a la información pública educativa

En septiembre de 2016 se sancionó en Argentina la Ley 27.275 de acceso a la información pública. Esta ley vino a cubrir un vacío en términos de derechos y procedimientos que regularan y garantizaran la transparencia en la gestión de la actividad estatal y de gobierno, creándose a partir de ella la Agencia de Acceso a la Información Pública, ente autárquico en el ámbito del Poder Ejecutivo de la Nación que garantiza el ejercicio del derecho a la información en el marco de dicha ley (Basterra, 2017).

Esta ley representó un hito en la legislación argentina y en las prácticas respecto de la información estatal en el país, en cuanto establece como principio gubernativo la “transparencia activa”, teniendo como objetivo la máxima accesibilidad a la información producida o en poder del estado.

Gracias a su implementación, cada área de gobierno posee responsables de facilitar respuestas a los pedidos de acceso a la información pública, siendo la Agencia de Acceso a la Información Pública el organismo que, en los casos de denegación, intercede para garantizar el derecho al acceso a la información pública.

Esta agencia, en los casos en que el sujeto obligado niega la información, solicita a ambas partes amplíen sus fundamentos, se expide y publica sus resoluciones. Estas son públicas, lo que permite consultarlas y trabajar sobre su contenido.

De un análisis de casos, es posible resumir que

la Secretaría de Evaluación e Información Educativa del Ministerio de Educación de la Nación niega indebidamente:

- pedidos de información sobre la matrícula (cuántos alumnos hay por escuela o por municipio);
- pedidos de información sobre la cantidad de secciones (cursos por escuela o por municipio);
- resultados de evaluaciones educativas desagregadas a nivel subprovincial.

La Agencia de Acceso a la Información Pública (AAIP) emitió, ante diferentes reclamos, resoluciones que ponen en evidencia cómo este comportamiento se encuentra reñido con el derecho vigente (AAIP, 2023). Resumiré a continuación algunas de estas consideraciones.

En las resoluciones de la AAIP se señala que la Secretaría de Evaluación e Información Educativa no puede meramente invocar “confidencialidad” sobre cualquier información utilizada con fines estadísticos. Para hacerlo, es preciso que o bien la información sea tomada a terceros exclusivamente con fines estadísticos (bajo promesa de mantener su reserva), o bien sea parte del plan anual de estadísticas y censos (artículos 4 y 5 de la Ley 17.622) o por indicación de la autoridad que coordina el Servicio Estadístico Nacional (artículos 1 y 3 del Decreto 3110/70) (AAIP, 2019).

Asimismo, al respecto de la confidencialidad que el Ministerio evoca en varias de sus denegaciones, la AAIP responde que no es comprensible que el Ministerio niegue pedidos de información sobre datos que al mismo tiempo difunde sin protección. En este sentido, si los resultados de una evaluación educativa fueran confidenciales, no sería adecuado que al mismo tiempo la SEIE dé acceso a ellos en las escuelas y en sus “comunidades educativas” (tal como lo hace) tanto al personal educativo a como sujetos ajenos a las instituciones, sin aplicar al hacerlo ninguna de las salvaguardas que legalmente caben a un dato confidencial en poder del Estado. No es claro a qué definición de “confidencial” alude el Ministerio al comportarse de esta manera, pero no es aquella que reconocen las regulaciones vigentes (es decir, según la legislación, si algo es confidencial no puede difundirse a ninguna escala pública).

Tanto con respecto a la confidencialidad de la información educativa como al artículo 97 de la Ley de Educación Nacional 26.206 sobre la estigmatiza-

ción al que la Secretaría de Evaluación e Información Educativa repetidamente refiere, la AAIP sostiene que no debe apelarse a elementos normativos aislados sin realizar una interpretación de conjunto de la normativa. En la resolución 2019/80 señala que para la correcta aplicación de las regulaciones vigentes debe realizarse en todos los casos una lectura armónica del cuerpo normativo, que considere en este caso la nueva legislación referida al acceso a la información pública (AAIP, 2019). En este sentido, expresa que la existencia de la Ley 27.275 obliga a una relectura de las demás normativas, de modo que las antiguas regulaciones, incluso si vigentes, no hagan abstracta la sanción de las nuevas disposiciones.

Por si esto no tuviera suficiente peso como criterio interpretativo general (la búsqueda de una adecuación armoniosa de la normativa preexistente), en la resolución 2019-80 también señala que la Corte Suprema de Justicia de la Nación estableció que ante conflictos respecto de la necesidad de dar difusión o mantener en reserva información pública, el criterio a seguir debe ser que “en la ponderación de los derechos que se encontrarían en conflicto (...) debe prevalecer el principio de máxima divulgación de la información pública” (AAIP, 2019, p. 6) (ver también Aveldaño et al., 2016).

Adicionalmente, sobre la cuestión de la problemática de la estigmatización de las entidades educativas, la AAIP indica que la mera especulación acerca de la estigmatización que la publicación de la información pueda provocar es insuficiente para justificar la restricción de la garantía de acceso a la información pública. Y en particular, expresa que “la experiencia demuestra que la estigmatización de las instituciones educativas ya existe en la actualidad y no es producto de políticas de transparencia, bien ausentes hasta ahora, sino de una marcada desigualdad social” (AAIP, 2019, p. 9). Agrega en este caso que:

La mera afirmación de que la publicidad de información referida al desempeño educativo a nivel de cada institución podría estigmatizarlas es insuficiente no sólo por lo abstracta y especulativa que resulta, sino además porque el daño que se procura evi-

tar se verifica como una grave problemática social que ya existe en la actualidad. (AAIP, 2019, p. 9)

A pesar de la fuerza de estos argumentos, cabe señalar que la AAIP se pronuncia siempre en el marco de pedidos individuales, por lo que las dependencias estatales —si bien podrían tomar las resoluciones previas de las AAIP como antecedentes relevantes— con frecuencia persisten en rechazar pedidos con argumentos largamente cuestionados por la Agencia de Acceso a la Información Pública.

Parece necesario aquí que se produzca un diálogo entre los agentes estatales que pueda gestar una conformidad de criterios entre el ente rector de la interpretación del acceso a los datos públicos (la AAIP) y, en este caso, la Secretaría de Evaluación e Información Educativa del Ministerio de Educación de la Nación. En este sentido, el Consejo Federal de Educación parece un espacio adecuado para este tipo de diálogo, a la luz de que existen posturas heterogéneas al interior del país. Como caso emblemático de este diálogo pendiente sobre apertura de datos cabe mencionar que la Provincia de Buenos Aires, la cual concentra cerca del 25% de los establecimientos educativos del país, publica en su Mapa Educativo la matrícula de alumnos por escuela, la misma que el Ministerio Nacional niega por afirmar que es confidencial.

Tal como se detalló anteriormente, este artículo busca cubrir dos aspectos de la publicación de datos educativos en Argentina: el grado de confidencialidad atribuido a la información, y la forma en que dicha confidencialidad se refuerza. Hasta aquí pudimos ver cómo, según el ente encargado de dirimir en cuestiones de apertura de datos, existe una sobre-denegación de datos educativos. A continuación, pondremos atención en cómo se ejerce esta sobreprotección de datos, para evaluar su efectividad y su rigurosidad de la misma.

Fiabilidad del esquema de protección de datos

Analizar el esquema de protección basado en la utilización de ‘identificadores enmascarados’ (o identificadores de fantasía, como los llama el Ministerio de Educación) supone evaluar en qué medida dichos

identificadores son producidos de un modo en que los potenciales usuarios de la información no pueden revelar la identidad de cada elemento (en este caso, establecimientos educativos).

En la documentación que acompaña la publicación de los datos del Ministerio de Educación, se indica: “Como se menciona en la presentación de este documento, si bien la información se presenta a nivel del establecimiento, la identificación del mismo se encuentra enmascarada con un valor ficticio” (DIEE, 2023a, p. 4)

En la medida en que no se indica de qué manera se enmascara esa información, emerge la preocupación sobre qué tan efectivo (en términos de seguridad) es el esquema seleccionado.

Para esto, analizaremos a partir de los listados de códigos únicos de establecimiento y de identificadores de fantasía qué tan posible es “deducir” el valor de uno a partir del otro. Las dos primeras hipótesis a considerar, son:

1. Que los “identificadores de fantasía” conformaran una lista de números aleatorios, por lo que existiría una tabla de correspondencia secreta entre dichos números aleatorios y los identificadores originales.

2. Que los “identificadores de fantasía” son el resultado de una transformación de los cueanexo originales.

Para validar estas hipótesis, se obtuvieron del sitio web del Ministerio una base de datos con el conjunto de establecimientos de la Ciudad de Buenos Aires con sus identificadores de fantasía, y —a partir del Padrón de Establecimientos— se obtuvo otro listado con el conjunto de escuelas de la Ciudad de Buenos Aires (con sus cueanexos). A partir de ambos listados, se procedió a realizar un análisis de sus características generales y de las relaciones entre ellos, que se presenta a continuación.

1. Análisis de las características de los cueanexo

Los “cueanexo” son identificadores compuestos por caracteres numéricos, de nueve cifras de longitud (Tabla 1). Al revisar el listado para diferentes zonas del país, pudo observarse que los dos primeros dígitos cueanexos representan a la jurisdicción (INDEC, 2012).

Los dos últimos dígitos del cueanexo son utilizados para identificar las diferentes sedes que pueda

tener un establecimiento. La sede principal tiene 00 al final, y de haber otras se identifican como 01, 02, etc.

2. Análisis de las características de los identificadores de fantasía

En el caso de los identificadores de fantasía, también se trata de elementos compuestos por caracteres exclusivamente numéricos (Tabla 1). La longitud de los mismos es de quince dígitos, y no es posible reconocer a primera vista la presencia de la jurisdicción en la formación del número. Al examinar el listado de identificadores de fantasía, se advierte la repetición de la parte inicial en su continuación. El primer Id listado tiene una primera parte '100502000', para luego repetirse '100502' hasta alcanzar los quince dígitos. De esta forma, en los Ids de fantasía la longitud fija de quince caracteres parece lograrse por la repetición de un mismo número de longitud más corta, hasta alcanzar los 15 dígitos.

Tabla 1.

Selección de cueanexos e identificadores de fantasía de la Ciudad de Buenos Aires (jurisdicción 02)

Cueanexos (9 dígitos)	Ids fantasía (15 dígitos)
020000100	100502000100502
020000200	100055200010005
020000300	100555200010055
020000500	100151520001001
020000700	100202020001002
020000800	100252520001002
020000900	100353520001003
020001000	100454520001004

Elab. propia en base a Ids de fantasía de base usuaria del Relevamiento Anual de Escuelas, CABA y Cueanexos provenientes del Mapa Educativo.

3. Análisis de frecuencias

Un paso elemental en un análisis criptográfico consiste en revisar la frecuencia con la que aparecen los signos en los mensajes. En nuestro caso, el análisis de frecuencias tiene la finalidad de ayudar a establecer qué tan probable es la hipótesis

de que los "Id de fantasía" se construyan en forma aleatoria, guardándose la tabla de correspondencia entre ellos y los cueanexos en forma privada. O que incluso, de no hacerse de este modo, evaluar si el enmascaramiento realizado utiliza un algoritmo seguro de hashing o de encriptación, los cuales garantizan una distribución homogénea de los valores resultantes (Sánchez et al., 2018; Bray, 2020).

Si observamos las frecuencias del primer conjunto de identificadores de establecimientos de la Ciudad de Buenos Aires (Tabla 2), vemos que la frecuencia del valor cero es muy superior a todos los demás dígitos, siendo 155 de los 330 caracteres (casi 47%). Asimismo, el número '1' muestra una frecuencia mayor que los demás dígitos (18,78%) haciendo improbable que este listado de números provenga de una producción aleatoria o de una codificación que utilice algoritmos conocidos de hashing o de encriptación.

Tabla 2.

Identificadores de Fantasía de la Ciudad de Buenos Aires (jurisdicción: 02)

Id fantasía	Frecuencias									
	0	1	2	3	4	5	6	7	8	9
100502000100502	9	2	2	0	0	2	0	0	0	0
100055200010005	9	2	1	0	0	3	0	0	0	0
100555200010055	7	2	1	0	0	5	0	0	0	0
100151520001001	7	5	1	0	0	2	0	0	0	0
100202020001002	9	2	4	0	0	0	0	0	0	0
100252520001002	7	2	4	0	0	2	0	0	0	0
100353520001003	7	2	1	3	0	2	0	0	0	0
100454520001004	7	2	1	0	3	2	0	0	0	0
100601020001006	9	3	1	0	0	0	2	0	0	0
100651520001006	7	3	1	0	0	2	2	0	0	0
100702020001007	9	2	2	0	0	0	0	2	0	0
100752520001007	7	2	2	0	0	2	0	2	0	0
101101020001011	7	7	1	0	0	0	0	0	0	0
101252520001012	5	4	4	0	0	2	0	0	0	0
100800048096100	8	2	0	0	1	0	1	0	2	1
100800048336100	7	2	0	2	1	0	1	0	2	0
100800048576100	7	2	0	0	1	1	1	1	2	0
100170012606310	6	4	1	1	0	0	2	1	0	0
100170012732310	5	4	2	2	0	0	0	2	0	0
100170012795310	5	4	1	1	0	1	0	2	0	1
100697002029091	6	2	2	0	0	0	1	1	0	3
100697002049291	5	2	2	0	1	0	1	1	0	3
Total	155	62	34	9	7	26	11	12	6	8

Elab. propia en base a identificadores de fantasía en base usuaria del Relevamiento Anual de Escuelas, CABA, 2021.

Este hecho permite suponer, tentativamente, que la secuencia de números de la que derivan los “Identificadores de fantasía” no constituye una tabla de números aleatorios creada para ese fin. Por esta razón, podemos provisoriamente aceptar que son un producto transformado del correspondiente cueanexo, y avanzar en la pregunta: ¿en qué consiste esa transformación?

4. Longitud efectiva del Identificador de fantasía

Al observar la tabla de identificadores de fantasía, se advierte —como se detalló anteriormente— que los identificadores de fantasía repiten su inicio hacia el final del número. La segunda columna de la tabla en la Tabla 3 muestra el resultado de “limpiar” de los “Id de fantasía” estos fragmentos repetidos, para evitar la parte redundante. En dicha tabla fueron clasificados los Id no redundantes según tuvieran pares de ceros en su final o en su parte media, regularidad que fue advertida al quitar la ‘cola’ o parte repetida de los Id de fantasía completos.

Tabla 3.

Identificadores de Fantasía de la Ciudad de Buenos Aires (Código de jurisdicción: 02).

Id fantasía*	Id fantasía sin "cola" **	Tipo de Id de fantasía
100502000100502	100502000	Con ceros al final
100055200010005	1000552000	
100555200010055	1005552000	
100151520001001	10015152000	
100202020001002	10020202000	
100252520001002	10025252000	Con ceros en el centro
100800048576100	100800048576	
100170012606310	1001700126063	
100170012732310	1001700127323	
100170012795310	1001700127953	
100697002029091	10069700202909	
100697002049291	10069700204929	

*resaltada la parte redundante del identificador

**resaltada la zona de ceros del identificador una vez removida la “cola”

Elab. propia en base a IDs de fantasía en base usuaria de cargos, CABA, 2021.

5. Patrón de crecimiento

Para proseguir, se elaboran explicaciones posibles respecto de la abundancia de números cero, así como también sobre la variación en el ancho de los dígitos.

Si nos dirigimos al padrón de establecimientos de la Ciudad de Buenos Aires, vemos que las escuelas de esa ciudad comienzan con los siguientes números de cueanexo: 020000100, 020000200, 020000300, 020000500, 020000700, 020000800, 020000900, 020001000, 020001100.

Esta lista, al igual que la lista revisada anteriormente, presenta números crecientes y números con ‘abundancia de ceros’. ¿Podrá deberse esta coincidencia a una transformación de números para obtener otros, transformación que mantiene su carácter creciente y su distribución de cifras?

En los casos incluidos en la tabla como ejemplo, los cueanexo crecen en sus cifras inferiores de “cue”: si descartamos los dos ceros de la derecha (que son los números de anexo), vemos que los cueanexo crecen linealmente a incrementos de uno: 0200001, 0200002, 0200003, etc. El crecimiento de los números de Id de fantasía, en cambio, resulta más acelerado que el crecimiento de los cueanexo: mientras que los cueanexo crecen de a una unidad, los Id de fantasía (Tabla 3) pasan en apenas doce números del orden de los millones al orden del billón.

Si mantenemos el supuesto de que uno es la transformación numérica del otro, pero que crecen a ritmos diferentes, cabría buscar identificar el tipo de transformación numérica utilizada.

6. Multiplicación directa

La multiplicación es una operación matemática que mantiene los ceros (cien por cualquier número entero, tendrá dos ceros a la derecha), y que de realizarse con un factor mayor a uno puede amplificar una tasa de crecimiento. Esto hace razonable evaluar si los identificadores de fantasía pueden ser el resultado de multiplicar los cueanexo. Para esto, se aplicaron diferentes estrategias para identificar si los identificadores de fantasía resultaban ser divisores exactos de ciertos cueanexo (o a la inversa) para diferentes localidades, no encontrándose patrones o coincidencias recurrentes en este sentido.

7. Multiplicación parcial

En la Tabla 3 fueron colocados casos testigo de los dos grandes grupos de identificadores de fantasía: aquellos que tenían ceros en su parte final (inicio del listado) y aquellos que tenían ceros en su centro (parte inferior del listado). En todos los casos, resultaba tan frecuente la presencia de dobles cero en los id de fantasía como lo era en los cueanexo (con ceros en su parte final).

La existencia de estos dos grupos (con ceros en el medio, y con ceros al final, a diferencia de los cueanexo que tienen ceros al final) permite especular, buscando otras relaciones, que como parte de las transformaciones el cueanexo sea separado en mitades e invertidas en su orden.

Por ejemplo, si tomamos 020001100, podríamos separarlo en mitades (02000 y 1100), multiplicar ambas por 2, teniendo como resultados 04000 y 2200, y luego juntarlas invirtiendo el orden: 22004000. Al hacerlo habría mantenido los ceros en su centro, modificado el número, y llegado a un valor que, de haber existido dígitos más grandes, habría crecido también en ancho.

Es interesante que estos pasos (multiplicar mitades e invertir el orden) produce el efecto que observamos antes, es decir, que algunos identificadores de fantasía sean números mucho mayores que otros. Si multiplicamos las mitades de un número (los primeros dígitos y luego los últimos) podemos arribar a este tipo de variabilidad, dado que las mitades (si separamos por una cantidad fija de dígitos) pueden tener órdenes de magnitud variables. Por ejemplo, el número 020001100, separado al medio, puede ser 02000 y 1100 (la segunda parte está en el orden de los miles). Pero el número 020000100 podría descomponerse en 02000 y 0100, es decir, 02000 y 100... teniendo 100 un orden menos de magnitud que 1100.

¿Cómo evaluar si esta clase de multiplicación de mitades tienen lugar en la construcción de identificadores de fantasía? De tratarse (la construcción del identificador de fantasía) de la concatenación de la multiplicación de dos mitades por un número igual, precisaríamos encontrar los divisores para ambas mitades. El procedimiento más simple sería utilizar el mismo divisor para ambas partes.

En la Tabla 4 podemos ver los números que teníamos en la Tabla 3 ("sin cola"), separados en dos partes y con ceros centrales para reconocer la

delimitación entre las mitades de los resultados (Parte 1 y Parte 2).

De las dos columnas, la Parte 1 es aquella que es candidata a dividirse manteniendo los ceros finales, por lo que vamos a suponer que corresponde a la mitad derecha del cueanexo. Al mismo tiempo, contamos con otra restricción, y es que estos divisores deben ser exactos (no dejar resto) y dejar la Parte 2 comenzando con el código de jurisdicción (02 en CABA). Que se trate de divisores que no produzcan resto en ninguna de ambas partes, y requerir de un 02 al inicio y un 00 al final, simplifica la búsqueda de posibles divisores.

En la Tabla 4 puede verse que, a diferencia de lo ocurrido con la división simple, la búsqueda de estos divisores arrojó valores compatibles con las restricciones establecidas³.

Tabla 4.

Análisis del Id de fantasía y sus divisores (selección de casos)

Id de fantasía sin cola	Parte 1	Parte 2	Divisor	Parte 1 dividida	Parte 2 dividida	Ambas partes unidas	Longitud obtenida
100800048576	1008000	48576	24	42000	02024	020242000	9
1001700126063	1001700	126063	63	15900	02001	020015900	9
1001700127323	1001700	127323	63	15900	02021	020215900	9
1001700127953	1001700	127953	63	15900	02031	020315900	9

Elab. propia en base a IDs de fantasía en base usuaria de cargos, CABA, 2021.

De este ejercicio surgen algunos emergentes de interés. Al crecer el número (id de fantasía sin cola) también fueron creciendo los divisores que cumplían con los requisitos (24, 63, 101). El resultado de aplicar esos divisores (partes 1 y 2 divididas) produjo números de 5 cifras, los que concatenados generarían números de 10 dígitos (uno más que el cueanexo). Sin embargo, al observarse la lista puede verse que en la totalidad de números evaluados (la Tabla 4 es solo una selección ilustrativa de ellos) los números inicial y final de las partes 1 y 2 divididas son siempre idénticos (en la primera fila la Parte 1 empieza con 4 y la Parte 2 divide termina con 4, y lo mismo en las demás filas). Esto sugiere que al producirse

las mitades a multiplicar, la cifra central de los 9 dígitos del cuanexo se utiliza en ambas partes, dejándolas con 5 dígitos.

8. Correlación del multiplicador con el final del cuanexo

Si se observa en la Tabla 4, se advierte otra regularidad que también fue estable aumentado el listado de casos: hay una correlación entre el divisor (24, 63, 101) y los dos últimos dígitos del cue (sin los 00 de anexo... 20, 59, 97). Uno y otro difieren siempre en 4 unidades: cuando los últimos dígitos del cuanexo forman '20' (como en 20244[20]00), el divisor es 24. Cuando es 59 (como en 20011[59]00), el divisor es 63. Y lo mismo pasa cuando es 97, que el divisor es 101.

Para verificar estas hipótesis, fueron seleccionados casos adicionales (Tabla 5), los cuales muestran la estabilidad de las reglas reconocidos para identificadores con ceros en el centro. A partir de un divisor que permita comenzar con 02, no dar resto y finalizar en ceros, se obtienen cifras con el formato de los cuanexo, y el divisor es cuatro números mayor que los dígitos finales. De este modo, para hacer el camino inverso (obtener el Id de fantasía a partir del cuanexo), si el cuanexo fuera 02013[12]00, deberían multiplicarse ambas mitades por 12+4, es decir, por 16.

Tabla 5

Análisis parcial del Id de fantasía con ceros centrales y sus divisores, segundo grupo

Id de fantasía sin cola	Parte 1	Parte 2	Divisor que		Id original teórico	Longitud obtenida	
			permite dejar	Parte 1			Parte 2
			la parte 2 en	dividida			dividida
			020				
16650030165	166500	30165	15	11100	2011	020111100	9
49920032208	499200	32208	16	31200	2013	020131200	9
53210034221	532100	34221	17	31300	2013	020131300	9
56520036234	565200	36234	18	31400	2013	020131400	9
69960044286	699600	44286	22	31800	2013	020131800	9
73370046299	733700	46299	23	31900	2013	020131900	9

Elab. propia a partir de base usuaria de cargos, 2021.

En estos nuevos casos, los divisores se encontraron entre 15 y 23, manteniéndose los valores

coherentes con lo esperado a partir de las observaciones previas. Los números de cue fueron de 11 a 19, manteniendo en todos los casos una distancia de 4 con el divisor.

Con el fin de contrastar la aplicabilidad de este esquema de cálculo, se realizó una validación más extensa sobre 200 Identificadores con ceros en el centro, y en todos los casos los resultados fueron consistentes con lo hallado a partir de los Identificadores seleccionados en las Tablas 4 y 5.

9. Identificadores con ceros al final

Ahora que tenemos una posible comprensión de cómo pudieron ser convertidos los identificadores que en sus resultados tenían ceros en el centro (segundo grupo de la Tabla 3), vamos revisar los identificadores que tenían un formato diferente (doble cero al final).

Al analizar varios de estos identificadores (por su rango, frecuencias y posibles divisores) parecería adecuado considerar la hipótesis de que los ceros finales se explicaran por haberse concatenado las mitades sin invertir su ubicación original. En estos casos, para hacer la separación en mitades, vamos a considerar que la magnitud de las cifras es similar a la tabla anterior, en la que las mitades con ceros eran más 'anchas' que las mitades de cifras sin ceros. Así, el número 1001700126063 tenía 7 dígitos en la primera parte (1001700) y solo 6 en la segunda (126063). Construimos así primeras partes que tienen 5 dígitos, y segundas partes que tienen entre 5 y 6 dígitos (Tabla 6).

Tabla 6.

Análisis del Id de fantasía y sus divisores para casos con ceros al final (selección de casos)

Id de fantasía sin cola	Parte 1	Parte 2	Divisor que permite dejar la parte 1 en 02	Parte 1 dividida	Parte 2 dividida	Ambas	Longitud obtenida
1005552000	10055	52000	5	02011	10400	020110400	9
10015152000	10015	152000	5	02003	30400	020030400	9
10020202000	10020	202000	5	02004	40400	020040400	9
10025252000	10025	252000	5	02005	50400	020050400	9
10035352000	10035	352000	5	02007	70400	020070400	9

Elaboración propia en base a IDs de fantasía en base usuaria de cargos, CABA, 2021.

Debemos ahora, igual que antes, buscar divisores que sin dejar resto puedan dividir la primera parte para que su resultado comience con 02, y la segunda parte termine con doble cero.

Al utilizar los números listados al inicio y buscar los divisores que cumplieran con las reglas mencionadas, el número encontrado fue 5 en todos los casos (Tabla 3, columna “divisor que permite dejar la parte 1 en 02”). Esto condujo a números de 9 dígitos de longitud (fusionando los números de final e inicio de ambas partes por ser idénticos) y finalizados en ceros (tal y como son mayoritariamente los cueanexo) (Tabla 6).

Sin embargo, cabe señalar que ahora la distancia entre el final del cue (valores 04 en todos los casos) y el divisor no fue de 4 unidades, sino de 1 ($5 - 4 = 1$). Es decir, en los casos de la Tabla 4 se concatenaron las mitades “cruzadas” multiplicando por el final del cueanexo más cuatro; en estos casos, los valores se concatenaron en forma directa y sumaron al cue una unidad para construir el multiplicador.

Para llevar el procedimiento a más casos, buscamos dentro de la Ciudad de Buenos Aires identificadores de fantasía que tuvieran ceros al final, pero que no comiencen con “100” (divisor esperado: 5). En la Tabla 7 se presenta una segunda selección de casos con ceros al final. En ellos, se ve mantenerse la relación de +1 entre el divisor y los últimos dígitos del cueanexo.

Tabla 7.

Análisis parcial de identificadores de fantasía y sus divisores para más casos, con ceros al final

Id de fantasía sin cola	Parte 1	Parte 2	Divisor que permite dejar la parte 1 en 02	Parte 1 dividida	Parte 2 dividida	Id original teórico	Longitud obtenida
200220000	2002	20000	1	2002	20000	020020000	9
400440200	4004	40200	2	2002	20100	020020100	9
800881200	8008	81200	4	2002	20300	020020300	9
16016165600	16016	165600	8	2002	20700	020020700	9
20020209000	20020	209000	10	2002	20900	020020900	9
22033341000	22033	341000	11	2003	31000	020031000	9

Fuente: Base usuaria de cargos, Relevamiento Anual de Escuelas 2021.

En síntesis, encontramos para este segundo formato valores de cueanexo de entre 0 y 10, que produjeron multiplicadores entre 1 y 11, y que dieron lugar a números armados a partir de concatenar la primera parte primero y la segunda parte después. Esto tiene una lógica similar a lo observado para los identificadores con ceros en el centro, diferenciándose la distancia entre los números finales del cueanexo (1 y 4) y en la necesidad de invertir o no las mitades antes de unirlos.

10. Verificación de resultados

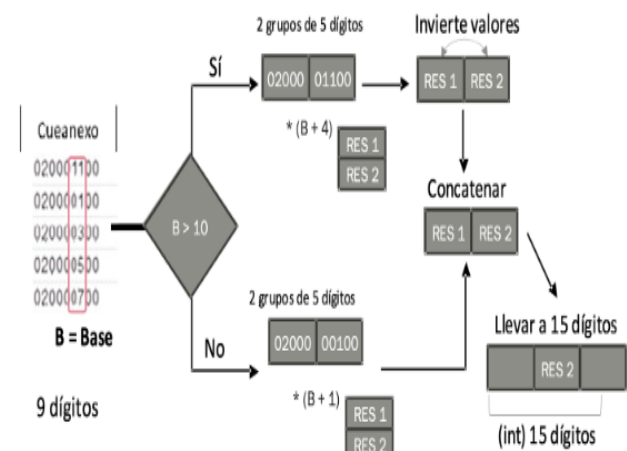
Con este conjunto de transformaciones posibles entre un identificador de fantasía y un cueanexo, resta realizar una verificación más exhaustiva para evaluar si este procedimiento aparea los elementos correctos. Es decir, no solamente se puede obtener un número compatible con la forma de un cueanexo y su jurisdicción a partir de un identificador de fantasía o viceversa, sino que el número obtenido sea en efecto el utilizado por el Ministerio de Educación para identificar cada establecimiento. Para poner esto en práctica, se implementó un procedimiento de cálculo de identificadores de fantasía a partir de cueanexos, el cual se detalla en la Figura 1.

La primera condición es que si la base (dígitos 6 y 7) forman un número mayor a 10, se utilizan los pasos de transformación descritos en la

parte superior del diagrama: separar en mitades y multiplicar cada mitad por la base más cuatro, e invertir el orden del resultado. En cambio, si la base es igual o inferior a 10, se multiplica las mitades por la base más uno. Finalmente, en ambos casos, se concatenan los resultados (RES 1 y RES 2) para formar el identificador. Finalmente, se los repite y se retiene los primeros 15 dígitos.

Figura 1

Diagrama de transformaciones y condiciones para la construcción de un identificador de fantasía (15 dígitos) a partir de un cueanexo (9 dígitos)



Elaboración propia a partir de análisis de identificadores.

La validación realizada del identificador de fantasía hipotético consistió en unir una base de datos de establecimientos con cueanexos con una base de datos de identificadores de fantasía a partir del identificador que fue calculado para cada establecimiento. Esto era útil debido a que hay atributos en común en ambas bases (como el sector [privado/público], el ámbito [rural/urbano] o la jurisdicción), los cuales debían coincidir en todos los casos si los identificadores producidos eran correctos.

En la Tabla 8 puede verse la clasificación para ocho provincias y las Ciudad de Buenos Aires de sector de los establecimientos según ambos listados luego de aparear las escuelas según el identificador de fantasía producido con el procedimiento de la Figura 1⁴.

En todos los casos (21.161 establecimientos) el sector atribuido por el procedimiento fue correcto, lo que parece satisfactorio para considerar que la asignación fue adecuada de los identificadores con los cueanexo.

Al realizar esto en el total de provincias, se puso de manifiesto la necesidad de una regla adicional, y era que a partir del cueanexo 10000 de la provincia 50 (es decir, a partir del valor 500010000) todos los casos eran tratados como si el final de su cue (la base) fuera mayor a diez (el procedimiento de la base más cuatro).

Tabla 8.

Escuelas de educación estatal del Relevamiento anual de escuelas 2018 apareadas con establecimientos del Mapa Educativo 2018 según reconstrucción de identificadores de fantasía (selección de jurisdicciones).

Jurisdicción	Sector (coincidentes ambas fuentes)		Sector (no coincidentes ambas fuentes)	
	Estatal	Privado	Estatal	Privado
Ciudad de Buenos Aires	602	3285	0	0
Buenos Aires	7242	703	0	0
Córdoba	2347	27	0	0
Catamarca	622	144	0	0
Chaco	1416	63	0	0
Chubut	353	475	0	0
Corrientes	1062	97	0	0
Entre Ríos	1516	311	0	0
Formosa	862	34	0	0
Total	16022	5139	0	0

Elab. propia a partir de Relevamiento anual de escuelas (2018) y Mapa Educativo 2018.

Incorporada esta condición, los resultados calculados para todas las jurisdicciones del país a partir de la información del padrón de establecimientos fue coincidente con la información disponible en el relevamiento anual de escuelas (por identificador de fantasía). Es decir, que luego de tomar el cueanexo de cada establecimiento del Mapa Educativo 2018, transformarlo para construir un identificador de fantasía, y comparar los sectores (estatal/privado) de cada escuela con el registro que tiene ese identificador en el Relevamiento anual de escuelas, el nivel de coincidencia fue exacto en el total de los establecimientos. Por tratarse de la coincidencia plena de decenas de miles de variables dicotómicas sucesivas, consideramos esto una prueba razonable de haber dado con el procedimiento original⁵.

Tomando esto en cuenta, para facilitar la replicabilidad del trabajo realizado se detalla en la Figura 2 la implementación en lenguaje Python de la función resultante para calcular los identificadores de fantasía a partir de los cuenaxo de establecimiento.

Figura 2.

*Implementación en Python del procedimiento de transformación de identificadores de establecimientos educativos**

```
def crear_id_fantasia(cueanexo):
    # 1. Separa en dos partes
    part1, part2 = cueanexo[:5], cueanexo[4:9]
    base = int(cueanexo[5:7])
    # 2. Multiplicador e inversión
    if base > 10 or int(cueanexo) >= 500010000:
        val_part2_trans = int(part1) * (base + 4)
        val_part1_trans = int(part2) * (base + 4)
    else:
        val_part1_trans = int(part1) * (base + 1)
        val_part2_trans = int(part2) * (base + 1)
    # 3. Une ambas partes
    result = f"{val_part1_trans}{val_part2_trans}"
    # 4. Replica para llegar a la longitud buscada
    ret = result + result + result
    # 5. Devuelve los primeros quince dígitos
    return ret[:15]
```

* Una versión funcional de este código puede consultarse en <https://github.com/discontinuos/fantasy-maker>.

Elaboración propia a partir de análisis de identificadores.

Conclusiones

Esta investigación tuvo como objetivo revisar el estado de situación respecto a la publicación de datos educativos por parte de la Secretaría de Evaluación e Información Educativa del Ministerio de Educación de la Nación. Repasemos los resultados a la luz de la información conocida y de las dimensiones analizadas.

Hemos visto que, según el corpus de resoluciones de la Agencia de Acceso a la Información Pública, el Ministerio de Educación podría dar cumplimiento a los lineamientos legales que son de carácter obligatorio en la gestión pública desde el año 2016, y difundir información sintética de cada escuela (como la cantidad de estudiantes, de docentes, de aulas) sin arriesgar su integridad ni la privacidad de sus alumnos. En virtud de lo dictaminado por la Agencia de Acceso a la Información Pública, sería deseable también poder difundir métricas referidas al rendimiento escolar en forma desagregada (por municipio, por escuela) a fin de resolver (y no acrecentar) la conocida situación de disparidad de los servicios educativos. Adicionalmente, cabe considerar si algunas de las 763 columnas que el ministerio distribuye como parte del relevamiento anual de escuelas pueda mantenerse en reserva, por no tener utilidad estadística o significar su difusión algún riesgo razonablemente justificado para los actores involucrados.

Lo que el Ministerio de Educación a través de la Secretaría de Evaluación e Información Educativa hace, en cambio, es distribuir la información sin filtros de columnas ni de filas (en el caso de las pruebas Aprender con nivel de detalle de una fila por alumno), sin indicación de a qué municipio o escuela pertenece cada registro. Al hacerlo, sostiene que la información es confidencial (incluido el atributo de municipio), y que se encuentra adecuadamente protegida para realizar totales a nivel provincial gracias a su identificador de fantasía.

Sin embargo, hemos mostrado que a partir de un análisis relativamente simple es posible replicar el esquema de cifrado utilizado. En consecuencia, se concluye que la Secretaría de Evaluación e Información Educativa expone información que afirma es confidencial, no siguiendo ningún estándar criptográfico ni de seguridad validado. En lugar de publicar variables relevantes a los especia-

listas en forma abierta y resguardar aquellas que podrían ser privadas, se publican todas sin una protección efectiva.

La información es un derecho, al igual que la privacidad, y las prácticas vigentes que de manera discrecional mantienen oculta información por técnicas de escasa confiabilidad logran el peor equilibrio entre accesibilidad y confidencialidad: los datos no están accesibles, y al mismo tiempo, no están seguros.

Debe ser una prioridad que los actores interesados en usar los datos para comprender mejor la realidad educativa los accedan sin restricciones arbitrarias, y que los responsables de producir y facilitar información tengan como prioridad abrir canales para su uso y difusión.

Escuchar propuestas sobre cómo podría disponibilizarse la información de maneras útiles y transparentes sería un primer paso en un camino que nos ayude a salir de la precariedad con que se organiza el acceso público a la información estatal educativa en Argentina. Propiciar que equipos de investigación académica, observatorios, oficinas estadísticas, puedan tomar la palabra y permitir que la oscuridad y la discrecionalidad dejen de ser la manera de facto de llevar la cosa pública en términos de difusión de datos y de acceso a información crucial y necesaria.

Los resultados obtenidos muestran que es de vital importancia que la información estadística sea producida y circulada a partir de estándares conocidos y de prácticas abiertas, donde la comunidad de usuarios sea protagonista de una transformación en el acceso a la información educativa del país.

Notas

¹ Esta investigación fue realizada con financiamiento del Consejo Nacional de Investigaciones Científicas y Técnicas de la Argentina (CONICET).

² La brecha fue aún mayor en el nivel secundario, donde 29,9% de los estudiantes en establecimientos de gestión privada en el último año de secundario lograron alcanzar niveles satisfactorios o avanzados en matemática, comparado a un 11,5% en establecimientos públicos (mostrando una diferencia en las chances relativas de más de dos veces y media) (DIEE, 2023c, pp. 89).

³ Para encontrar los divisores se construyó una tabla con los divisores posibles con ayuda de la herramienta Find all divisor of a number (disponible en <https://onlinemathtools.com/find-all-divisors>), manteniendo solo los que daban por resultado números iniciados en 02. Se omite el detalle de esa búsqueda para simplificar la presentación de resultados.

⁴ Para las jurisdicciones de Buenos Aires y Córdoba se encontraron 5 y 1 casos respectivamente en que el sector no era coincidente debido a errores en la información del Mapa Educativo, el cual publicaba como pertenecientes al sector público a establecimientos que, al investigarlos a partir de su nombre y su domicilio, eran de gestión privada, por lo que se corrigió el sector en estos 6 casos y se realizó nuevamente el conteo de coincidencias.

⁵ Adicionalmente al sector, se verificó con éxito también el ámbito (urbano/rural), para hacer más robusta la validación, control que no se incluyen aquí por razones de espacio.

Referencias

- Agencia de Acceso a la Información Pública (2019). Resolución 2019/80 de la Agencia de Acceso a la Información Pública. Reclamo Magioncalda al Ministerio de Educación de la Nación. <https://www.argentina.gob.ar/sites/default/files/resol-2019-80-apn-aaip.pdf>
- Agencia de Acceso a la Información Pública (2023). Resoluciones de la Agencia de Acceso a la Información Pública (RESOL-2021-230-APN-DNPDPAAP, RESOL-2022-61-APN-AAIP, RS-2018-13-APN-AAIP). <https://www.argentina.gob.ar/aaip/buscador-normativa>
- Aveldaño, A, Morales, A. y Ale, A. (2016). El derecho de acceso a la información pública en Argentina. Presentado en XVII Congreso Nacional y VII Congreso Latinoamericano de Sociología Jurídica. Tucumán (Argentina) (2016, Octubre 5-7).
- Basterra, M. I. (2017). La ley 27.275 de acceso a la información pública. Una deuda saldada. *Revista de derecho público*, 1, 11-42.
- Bonelli, S. (2021). Sistemas nominales de información educativa. *Argentinos por la Educación*.
- Bray, S. W. (2020). *Implementing cryptography using Python*. Indianapolis: John Wiley & Sons.
- Chávez, H. E. G., Novoa, A. E. S. y Huanca, M. M. (2022). Comparación de los derechos educativos en México, Colombia y Argentina. En S. Sevilla-Vallejo y R. Castrillo Soto (eds.) *Desafíos y retos de la educación en tiempos de pandemia: Aproximaciones educativas desde la ética, la lingüística y la literatura* (pp. 153-172).
- De Grande, P. (2019). El problema, ¿está en la casa? Relaciones personales y exclusión educativa en la Ciudad de Buenos Aires (2001-2010). *Páginas de Educación*, 12, 49-74.
- Dirección de Información y Estadística Educativa (2023a). *Bases Usuaras Relevamiento Anual*. Documento Metodológico. Ministerio de Educación de la ciudad de Buenos Aires.
- Dirección de Información y Estadística Educativa (2023b). *Aprender 2022. Educación primaria*. Informe nacional de resultados. Ministerio de Educación de la ciudad de Buenos Aires.
- Dirección de Información y Estadística Educativa (2023c). *Aprender 2022. Educación secundaria*. Informe nacional de resultados. Ministerio de Educación de la ciudad de Buenos Aires.
- El cronista (2023, Enero 27). El salario inicial docente es la mitad del mínimo para no ser pobre. <https://www.cronista.com/economia-politica/el-salario-inicial-docente-es-la-mitad-del-minimo-para-no-ser-pobre/>.
- Esquivel-Quirós, L. G., Barrantes, E. G. y Darlington, F. E. (2019). Marco de medición de la privacidad. *Revista Ibérica de Sistemas e Tecnologías de Informação*, (31), 66-81.
- Fung, B. C., Wang, K., Chen, R. y Yu, P. S. (2010). Privacy-preserving data publishing. *ACM Computing Surveys*, 42(4), 1-53.
- Guadagni, A. A. (2018). Los salarios docentes en el mundo y en Argentina. *Boletines del Centro de Estudios de la Educación Argentina*, 69, 3-11.
- Instituto de Investigaciones en Gestión, Desarrollo y Control de Organizaciones (2023). *Encuesta nacional: percepciones de la Educación Argentina*. Buenos Aires
- Instituto Nacional de Estadísticas y Censos (2012). *Códigos de las unidades geográficas utilizadas en el Censo Nacional de Población, Hogares y Viviendas 2010*. Buenos Aires: INDEC.
- Samarati, P. y Sweeney, L. (1998). Protecting Privacy when Disclosing Information: k-Anonymity and its Enforcement Through Generalization and Suppression. In, *Proceedings of the IEEE Symposium on Research in Security and Privacy* (pp. 384-393).
- Sánchez, S., Domínguez, P. y Velásquez, L. (2018). *Hashing: Técnicas y Hash para la Protección de Datos*. Documento de trabajo, Universidad Tecnológica de Panamá.
- Suasnábar, J.M. y Valencia, D. (2022). El uso de información estadística gubernamental para caracterizar desigualdades sociales y educativas en los territorios. *Propuesta Educativa*, 31(57), 58-74.
- Tiramonti, G. (2019). La escuela media argentina: el devenir de una crisis. *Propuesta Educativa*, 51, 78-92.

ISSN: 2362-3349

Cita sugerida: de Grande, P. (2025).
Problemáticas en la difusión de datos educativos.
El tratamiento de su confidencialidad en
Argentina. *Revista de la Escuela de Ciencias de
la Educación*, 1 (20), 100-117.

Recibido: 22 de febrero de 2024
Aprobado: 26 de junio de 2024
Publicado: 1 de enero de 2025
Facultad de Humanidades y Artes - UNR