
Policy Brief #5

Uso estratégico de datos e inteligencia artificial en la educación

**Gobierno digital
e innovación pública**

Uso estratégico de datos e inteligencia artificial en la educación

Tiempo de lectura: 10 minutos

Temas asociados: Deserción escolar - Educación – Innovación pública - Inteligencia artificial

Eje: gobierno digital e innovación pública

Nota de la Dirección de Transformación Digital de CAF -banco de desarrollo de América Latina

Esta nota está basada en el estudio “Uso estratégico de datos e inteligencia artificial en la educación” elaborado por Cecilia María Vélez White

Elaboración de la nota: Cecilia María Vélez White, consultora de CAF

Nota supervisada por: María Isabel Mejía, responsable de la agenda de Gobierno Digital e Innovación Pública en CAF

Revisión y comentarios: Claudia Flores, Nathalie Gerbasi, María Isabel Mejía, Martha Rodríguez y Antonio Silveira

© 2022 Corporación Andina de Fomento

Las ideas y planteamientos contenidos en esta nota son de exclusiva responsabilidad de su autor y no comprometen la posición oficial de CAF.

Resumen

- Este *policy brief* es el quinto de una serie de documentos de política sobre el uso e impacto de la inteligencia artificial en el sector público. Estos documentos sintetizan los principales hallazgos, conclusiones y recomendaciones de **ExperienciaIA**, el reporte regional lanzado por CAF en septiembre de 2021, con el objetivo de sensibilizar sobre las oportunidades y los retos del uso estratégico de datos e inteligencia artificial (IA) en el sector público en la región.
- Esta investigación se realizó con el fin de documentar casos de implementación de proyectos en el uso de datos e inteligencia artificial (IA) en el sector educativo, para el diseño, implementación y seguimiento de políticas públicas. Esto con el fin de *llegar* a recomendaciones sobre la posible utilización de estas herramientas por las administraciones del sector educativo en América Latina, para mejorar su eficiencia y eficacia.
- Se hace una revisión del estado del arte sobre el uso de datos y de la IA en el sector educativo, encontrando que en las últimas décadas se ha venido utilizando la analítica como una herramienta para personalizar la enseñanza. Esta herramienta, al generar mayor conocimiento y seguimiento de cada estudiante posibilita el diseño de soluciones personalizadas de aprendizaje. Las soluciones utilizadas se han focalizado en la generación de alertas tempranas y en la identificación de las intervenciones, pero aún no plantean las acciones que deben ser emprendidas para mejorar el proceso de enseñanza aprendizaje. Adicionalmente, por el corto espacio de tiempo de su aplicación, existe en general dificultad para la evaluación formal de las experiencias en función de su impacto real en el aprendizaje o incluso para evaluar si cumplieron su objetivo.
- Asimismo, se describen las limitaciones que por el momento tienen los sistemas de detección temprana de riesgo de deserción desarrollados, relacionadas con el funcionamiento, el seguimiento de las trayectorias mas allá del colegio, y la posibilidad de usar modelos mas sofisticados que permitan identificar subgrupos específicos.
- Los casos seleccionados ilustran el uso de datos y de modelos de detección temprana de riesgo de deserción en tres geografías diferentes: Estados Unidos (Wisconsin), Australia (Victoria), y Argentina (Provincia de Buenos Aires, PBA). Los criterios de selección iniciales tuvieron que ver con la sostenibilidad del proyecto y sus resultados comprobables, criterios que cumplía el desarrollo del sistema de información de Wisconsin (WISE) y su Drop Out Early Warning System (DEWS). Sin embargo, en el desarrollo de la investigación se encontró importante el análisis de los otros dos casos: el Student Mapping Tool (SMT) de Victoria, por la simplicidad de la herramienta, y el proyecto Asistire de la PBA, por tratarse de una experiencia en América Latina.

Caso Estados Unidos

El estudio hace una descripción detallada del proyecto del estado de Wisconsin, por cuanto es el que mayor permanencia ha tenido en el tiempo (desde el 2012), es el mas comprensivo, presenta resultados de impacto en las variables que pretendía intervenir, y supera algunas de las dificultades que se habían detectado en el uso de los modelos de detección temprana de deserción.

El Wisconsin Information System for Education (WISE) es un proyecto que busca ofrecer información sobre el sector educativo tanto al Gobierno Federal como a los distritos y escuelas del estado de Wisconsin, y al público en general. Tiene como objetivos una mayor transparencia en la gestión del sector, mayor capacidad de los ciudadanos para realizar la elección de los colegios de sus hijos y hacer seguimiento a las políticas públicas, y lograr que las escuelas del Estado usen los datos para la planeación y para el mejoramiento continuo.

Los componentes del WISE son:

- El WISEdata (recolección y depuración) y el WISElearn (procesos para hacer la información comprensible).
- El portal WISEdash que da acceso a la información a los diferentes niveles de usuarios –(Gobierno Federal, ciudadanos, distritos y escuelas).
- El WISEexplorer (proyecto de capacitación y uso).
- El WISEgrants, un sistema del acceso a subsidios del Gobierno Federal que se viene desarrollando desde el 2010 con el liderazgo del Departamento de Instrucción Pública (DIP) del Estado, y con cofinanciación del Gobierno Federal; entró en funcionamiento en el 2012.

El proyecto fue implementado por un equipo interdisciplinario del DIP e incluyó las siguientes actividades:

- Organización y depuración de los datos a partir del sistema de información existente en el estado (Sistema WINS). Al día de hoy la recolección de los datos se realiza en línea, lo que garantiza su permanente actualización y oportunidad.
- Estandarización de los criterios de recolección y clasificación de los datos.
- Desarrollo de la plataforma que permite la consulta de los datos y que ha sido especialmente flexible para garantizar el mejoramiento continuo en función de las demandas de los usuarios.
- Diseño e implementación del plan de capacitación de los distritos y de las escuelas, con el fin de que integren los datos a su planeación para garantizar el mejoramiento.
- Plan de comunicaciones sobre el sistema.

El uso del WISE ha venido creciendo. Las siguientes funciones básicas son utilizadas por todas las escuelas públicas y los distritos del Estado: la recolección de la información que se realiza sobre la plataforma en línea y el trámite de las solicitudes de recursos al Gobierno Federal que se hace asimismo a través del sistema. El programa WISEexplorer, con sus actividades de capacitación, amplía cada vez más la utilización de los datos en la planeación de las escuelas.

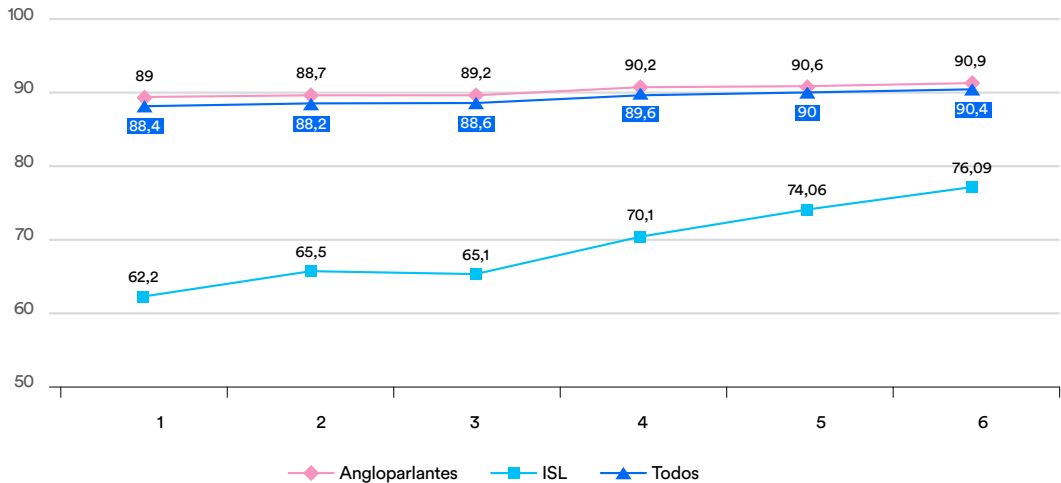
El *Drop Out Early Warning System* (DEWS) hace parte del WISE, y fue desarrollado en el 2012 para enfocar las acciones dirigidas a mejorar la oportunidad de graduación de los estudiantes, para garantizar sus trayectorias educativas y su tránsito a la educación superior. Especialmente se buscaba reducir el retraso en los tiempos de graduación de los alumnos que tienen el inglés como segunda lengua, un grupo poblacional creciente en el estado, así como el de los afrodescendientes. El modelo identifica en los grados 9º y 10º a los estudiantes en riesgo de quedar rezagados y convertirse en posibles desertores. Se realiza esa identificación mediante un sistema complejo de modelos predictivos, utilizando algoritmos y escogiendo entre estos los que arrojan resultados más ajustados a los datos. De esta forma se obtienen predicciones

más confiables sobre los estudiantes que los que se obtendrían con el uso de regresiones. Los indicadores utilizados en el modelo son: asistencia, disciplina, movilidad, resultados de las pruebas del Estado y datos demográficos. El modelo se corre anualmente para cada cohorte y arroja los resultados para cada estudiante.

La capacitación en el uso del modelo está integrada a la del sistema WISE en general y la realiza el

equipo del WISExplorer. El impacto del proyecto se pudo evidenciar por el aumento en el porcentaje de los estudiantes que se gradúan a tiempo, impulsado principalmente por la reducción de la brecha entre los estudiantes con el inglés como segunda lengua y los angloparlantes. Este seguimiento se puede realizar a través del WISEdash (ver Gráfico 1).

Gráfico 1. Porcentaje de estudiantes angloparlantes y con inglés como segunda lengua (ISL) de la cohorte que se gradúa a tiempo



Fuente: Tablero de datos WISE (Wisconsin Department of Public Instruction, 2021)

Caso Australia

El segundo caso analizado es el *Student Mapping Tool* (SMT) que se desarrolló en el Estado de Victoria (Australia) en el 2004. Este resultó de la iniciativa del Departamento de Educación del Estado en el marco de su política de reducir la deserción. Contrataron a un exrector para desarrollar la herramienta para identificar el nivel de riesgo de deserción de los estudiantes y lograr que las instituciones desarrollarán acciones para prevenirla. El modelo funciona mediante una hoja de Excel en la que se presentan, para todos los estudiantes, las variables determinantes de la deserción, unas existentes en el Sistema de Información Estatal, y otras que debían adicionar

las instituciones educativas. Las instituciones deben, en columnas adicionales, establecer la calificación del riesgo de los estudiantes, registrar las acciones para apoyar a los alumnos identificados y hacer el seguimiento de estas.

Las variables cargadas automáticamente son: asistencia (ausencias injustificadas y porcentaje de días atendidos); resultados en lecto-escritura y matemáticas (la evaluación de los profesores, y los resultados de los tests del Estado estandarizados); número de suspensiones internas de las clases, o externas por fuera de la escuela); ocupación de los padres, clasificada de acuerdo a sus condiciones económicas; estatus de la vivienda, considerando a quienes viven con los padres, con uno de ellos, o en hogares especiales; y grupos con necesidades especiales como

discapacitados, refugiados, etc. Las instituciones, de acuerdo con las particularidades de su contexto, pueden adicionar variables recogidas a través de encuestas o entrevistas. Las variables en las que se consignan el nivel de riesgo deben ser diligenciadas por la institución educativa, agregando la información de las primeras columnas, en las que aparecen señaladas para cada variable los estudiantes en riesgo. Para cada uno de los estudiantes que se identifican en riesgo se señalan las acciones que se desarrollan como planes individuales en el curso de las clases, o refuerzo mediante clases adicionales de lengua o de matemáticas, así como la participación en programas especiales del Estado o en programas de bienestar de la institución.

Se establecieron como responsables de la implementación de la herramienta en las escuelas los coordinadores y se realizó un amplio programa de difusión de la herramienta, lo que generó mucho entusiasmo inicial entre los rectores. Sin embargo, no se generaron incentivos, ni se estableció un plan de seguimiento. Dado el nivel de la autonomía de las escuelas en Australia, muy pocas implementaron el proyecto, y adicionalmente, con el cambio de administración cambiaron las prioridades. En la actualidad la herramienta sigue disponible en la página Web del Departamento de Educación, pero no se ha vuelto a actualizar, salvo la que el constructor de la herramienta realizó en el 2015.

Caso Argentina

El tercer caso, el modelo de identificación de estudiantes en riesgo de abandono, se desarrolló en la Provincia de Buenos Aires (PBA), Argentina, en el marco del proyecto Asistiré financiado por el Banco Mundial. El proyecto inició en el 2017 y consistía en la recolección de las estadísticas de asistencia mediante tabletas que se distribuían en las escuelas, la asesoría a estas por parte de un equipo interdisciplinario para desarrollar acciones que permitieran prevenir la deserción y la coordinación

en una mesa de trabajo acerca de las acciones de inclusión de diversas entidades estatales.

El modelo se contrató con la Universidad de Buenos Aires (UBA) con el fin de identificar estudiantes en riesgo. Trabajaron con los datos de dos años 2017-2018, dada la precariedad de los datos anteriores en la Provincia. Utilizaron el programa CaBoost con el fin de realizar las predicciones utilizando la metodología de boosting, cuyo funcionamiento se basa en árboles de decisión en los que se analizan las variables predictoras. Se llegó a un razonable nivel de precisión del sistema.

Hallazgos

En el desarrollo de la investigación se evidenciaron elementos que hacen posible el buen uso de los datos y la IA en el sector educativo y que su utilización genere impactos en la mejora de las trayectorias educativas de los estudiantes. Entre estos están: la apropiación del proyecto por parte de las instancias administrativas del sector, la calidad técnica y flexibilidad de las herramientas utilizadas, la planeación y la financiación de largo plazo tanto del desarrollo como de la operación del sistema, la comunicación sobre la implementación del proyecto y la capacitación en el uso del instrumento, el diseño de procesos de mejoramiento permanente del sistema, los incentivos para su uso, y los procesos de evaluación y seguimiento. Adicionalmente, resulta muy útil poner a disposición del público en general la información con el fin de que además de lograr una mayor participación en el proceso educativo, se conviertan en veedores del funcionamiento de los sistemas.

Asimismo, se encontraron algunas características que hacen fracasar la implementación de estos sistemas. Entre ellas:

- El desarrollo del proyecto por externos a la administración o por un área especializada sin contacto con las demás, el éxito depende de la interdisciplinariedad de los equipos.

- La inexistencia de una clara definición de las responsabilidades de todos los niveles de la administración en el diseño y operación del proceso, especialmente de las instancias que asesoran a las escuelas, acompañada del desarrollo de las capacidades para hacerlo.
- La inflexibilidad de los sistemas para adaptarse a cambios de políticas, lo que los hace vulnerables en los cambios de gobierno.

Recomendaciones

En cuanto a la aplicabilidad del uso de estas herramientas en América Latina se encontró que están dadas las condiciones para su desarrollo. En primer lugar, el mejoramiento de los sistemas de información que se viene dando en el sector en la región, que a su vez puede ser apoyada por el uso de los datos en el diseño e implementación de la política pública. En la actualidad se da una alta sub-utilización de los datos consignados en las bases de datos de los ministerios y allí se encuentran las variables claves para la descripción de muchas de las condiciones de los alumnos, lo que permitiría ofrecer a las instituciones educativas informaciones de vuelta (porque ellas son las que originan la información) muy útil para personalizar la enseñanza y mejorar los aprendizajes.

En general los ministerios han venido desarrollando sus capacidades técnicas y tendrían las suficientes para desarrollar bien sea una herramienta simple como el Student Mapping System de Australia, o unas mas complejas como el DEWS de Wisconsin. El desarrollo del modelo en la PBA nos demuestra que con series mas re-

ducidas en el tiempo es posible desarrollar modelos con buenos niveles de confiabilidad.

Conclusiones

Teniendo en cuenta las características de las administraciones públicas en América Latina es importante tener en cuenta algunas conclusiones de la investigación:

- Los modelos deben estar articulados a la gestión de los sistemas de información, y es especialmente importante llegar al nivel de las instituciones educativas, ya que son estas las que pueden generar impacto con su uso. Para disminuir el riesgo que representan los cambios de Gobierno en la implementación de los proyectos, es importante que los equipos que permanecen en las administraciones participen en la elaboración de las herramientas. Adicionalmente, enmarcar el uso de estos modelos en la finalidad mas amplia del mejoramiento de los sistemas de información, permite que se fortalezca en general la capacidad del uso de los datos en las administraciones y se puedan adaptar los instrumentos para que sean utilizados en cualquier tipo de política.
- La participación de agentes externos y de la academia puede ser necesaria para garantizar el correcto desarrollo de los aspectos técnicos. Pero es fundamental que el manejo del proyecto y los equipos ejecutores las administraciones tengan el liderazgo. Sea como asesores o como contratistas de algunos módulos, los externos deben responder a las prioridades y necesidades de la administración.

- En la financiación de los proyectos es necesario prever los costos de inversión necesarios para mejorar los sistemas y poner a andar los modelos, pero es igualmente importante prever los costos recurrentes que la operación de dichos sistemas y de las herramientas los generan. Esto es fundamental porque muchos de estos proyectos terminan archivados por la ausencia de operadores permanentes.
- En América Latina ha existido una ausencia de demandas ilustradas por parte de la sociedad hacia el sector educativo. En general

se hacen críticas permanentes y generales al sistema, pero paradójicamente la mayoría de los padres de familia creen que la educación en el colegio de sus hijos es buena. Las ventajas de generar información pública accesible y flexible, además de vincular al sistema educativo de manera mas estrecha a los padres de familia y a ciudadanos en general, se traduce en una mayor dinámica de mejoramiento del sector. La comunicación del proyecto es un componente esencial que, además de facilitar la ejecución, juega a favor de la permanencia del proyecto.

Bibliografía

Bagaskorowati, R. (2020). «Puspitaningrum Rini and Gunadi Asep. The development of student mapping tools application (SMT) as an early detection system for student-at-risk». *International Journal of Psychosocial Rehabilitation*, vol. 24, n.o 1, enero de 2020.

Bianchi, B., Pietto, M. L. y Kamienkowski, J. E. (2019). «Estimación de la interrupción de las trayectorias escolares en escuelas secundarias públicas de la provincia de Buenos Aires». Programa Manos en la DATA. Universidad de Buenos Aires.

Bowers, A. J., Sprott, R. y Taff, S. (2010). «Do we know who will drop out? A review of the predictors of dropping out of high school: precision, sensitivity and specificity». *Columbia University Libraries* [en línea]. 2013 Articles. <https://doi.org/10.1353/hcj.2013.0000>.

CAF (2018). «Políticas para promover la culminación de la educación media en América Latina y el Caribe. Lecciones desde México y Chile». CAF [en línea].

Ferguson, A., Brasher, A., Clow, D., Cooper, A., Hillaire, G., Mittelmeier, J., Rienties, B., Ullmann, T. y Vuorikari, R. (2016). *Research evidence on the use of learning analytics, implications for education policy*. Joint Research Centre Science for Policy Report. DOI:10.2791/955210.

Fiszbein, A. y Stanton, S. (2018). *Education in Latin America and the Caribbean: possibilities for United States investment and engagement*. Washington, D.C.: Inter-American Dialogue.

Knowles, J. E. (2015). «Of needles and haystacks: building an accurate statewide dropout early warning system in Wisconsin». *Journal of Educational Data Mining*, vol. 7, n.o 3.

Macfadyen, L., Dawson, S., Pardo, A. y Dragan, G. (2014). «Embracing big data in complex educational systems: the learning analytics imperative and the policy challenge». *Research & Practice in Assessment*, vol. 9, pp. 17-28.

National Dropout Prevention Center (2020). *Effectives strategies for dropout prevention*.

National Forum on Education Statistics (2018). *Forum guide to early warning systems* (NFES2019035). U.S. Department of Education. Washington, DC: National Center for Education Statistics. <https://nces.ed.gov/pubs2019/NFES2019035.pdf>.

Papamitsiou, Z. y Economides, A. A. (2014). «Learning analytics and educational data mining in practice: a systematic literature review of empirical evidence». *Educational Technology & Society*, 2014. 17(4): pp. 49-64.

Román, M. (2013). «Factores asociados al abandono y la deserción escolar en América Latina: una mirada en conjunto». *Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, vol. 11, n.o 2, pp. 33- 59. Madrid, España: RINACE. <https://www.redalyc.org/pdf/551/55127024002.pdf>.

Student Mapping Tool, <https://futureconnect.org.au/student-engagement-retention/student-mapping-tool/>

Sunbok Lee, and Jae Young Chung, *The Machine Learning-Based Dropout Early Warning System for Improving the Performance of Dropout Prediction*, *Applied Sciences review* (2019).





caf.com
@AgendaCAF
transformaciondigital@caf.com