FINAL TECHNICAL REPORT_INICIATIVA LATINOAMERICANA POR LOS DATOS ABIERTOS

Carla Bonina , Martin Harraca , Lorena Etcheverry , Neolia Beltramelli , Romina Colman , Anna Spinardi , Bárbara Battistotti , Zinnya del Villar

Carla Bonina , Martin Harraca , Lorena Etcheverry , Neolia Beltramelli , Romina Colman , Anna Spinardi , Bárbara Battistotti , Zinnya del Villar

©2024, CARLA BONINA , MARTIN HARRACA , LORENA ETCHEVERRY , NEOLIA BELTRAMELLI , ROMINA COLMAN , ANNA SPINARDI , BÁRBARA BATTISTOTTI , ZINNYA DEL VILLAR

This work is licensed under the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/legalcode), which permits unrestricted use, distribution, and reproduction, provided the original work is properly credited. Cette œuvre est mise à disposition selon les termes de la licence Creative Commons Attribution (https://creativecommons.org/licenses/by/4.0/legalcode), qui permet l'utilisation, la distribution et la reproduction sans restriction, pourvu que le mérite de la création originale soit adéquatement reconnu.

IDRC GRANT / SUBVENTION DU CRDI : - INNOVATIONS TO MONITOR KNOWLEDGE-SHARING RESULTS – EXPLORING THE OVERTON PLATFORM TO TRACK POLICY INFLUENCE

Tracking Research influence in Education Policy: Early Warning Systems in Latin America

AUTHORS:
NOELIA BERTRAMELLI,
LORENA ETCHEVERRY
AND CARLA BONINA

IDRC Project Number 109972-001

Innovations to monitor knowledge sharing results – Exploring the Overton platform to track policy influence

Tracking Research influence in Education Policy: Early Warning Systems in Latin America

Authors: Noelia Beltramelli, Lorena Etcheverry and Carla Bonina

31 May 2024

Abstract

We employ the Overton tool to explore the evidence regarding Early Warning Systems being implemented in education in the public sector in Latin America. Both policy documents and cited scholarly articles and their main features were considered in the analysis. We found that the geographical distribution of policy documents in Overton aligns with the literature on the subject, which explains the presence of these systems for nearly fifteen years in the United States and Europe, along with their emerging application in Latin American countries (Perusia and Cardini, 2021; Bowers, 2021). Evidence of policy documents from Latin America on this topic is scarce, mainly from Intergovernmental Organisations. Documents sourced from countries' governments' think tanks and others are less frequent, making it difficult to find detailed information about implementing these systems in each context. Moreover, we found that for these searches, it was crucial to contextualise the topic for different territories. We noticed that terminology varied in various countries around the same theme.

The underpinning research cited by the policy documents comes mainly from the Global North regarding authors' affiliation institutions. Institutions from the Global South, specifically Latin America, account for a small minority of the cited research. Most cited articles were published in journals from the Social Sciences, notably about Economics and Econometrics. The gender analysis shows that only one-third of the cited authors are women, and only one of the seven authors with the most citations is a woman.

We conclude from the research on this topic that Overton is a powerful tool for tracking policy documents and finding connections between topics. Although results were limited for this specific topic, it might be more beneficial for more general searches, such as tracking the use of AI in the public sector or the evolution of educational policies in Latin America.

Table of contents

1. Introduction

Education dropout rates in Latin America—the percentage of students who leave school before completing a particular level of education—are the focus of academic research and public policy. Although different countries exhibit a variety of situations depending on educational levels and related variables, upper secondary education (ages 15-17) dropout rates remain persistently high in the region, despite favorable macroeconomic conditions (Kattan and Székely, 2015). Furthermore, educational exclusion rates are exceptionally high at the upper secondary level for young people from lower socioeconomic backgrounds (Acevedo et al., 2020).[1]

Adding to this preexisting situation, the COVID-19 pandemic exacerbated the problem, with Latin America experiencing the longest interruption of in-person schooling internationally (Huepe, Palma, and Trucco, 2022). This occurred in a context of uneven distribution of digital equipment and connectivity, making it challenging to adapt to a distance education model (Almeyda et al., 2022). The educational costs of the pandemic in the region are estimated to set back years of progress in reducing exclusion, potentially raising exclusion levels above those seen before the pandemic (Acevedo et al., 2020).

In 2020, with schools partially or fully closing down, there was an urgent need for quality information to track students' pathways remotely, plan the return to in-person learning, and measure the impacts of the pandemic on student exclusion or the risk of dropping out. This situation put pressure on the region's educational management information systems (EMIS) to advance indicators to diagnose and address each country's situation effectively.

Some Latin American countries have already developed **early warning systems** (EWS) to identify students at risk of dropping out, but in 2020 and 2021, this trend has increased (UNESCO, 2022). These systems are an alternative that allows automating the identification of students at risk of dropping out based on academic, behavioral, socioeconomic, and other variables. However, it is essential to emphasize that these systems alone are not sufficient to effectively address the problem of school dropout. Predictive systems must be accompanied by tools that are adaptable to the specific context in which they are applied, as well as by efficient intervention plans to help retain students within the educational system (UNICEF, 2018; Arias Ortiz et al., 2021).

[1] The completion rate for primary education in the region was 93.4% in 2023. For lower secondary education the percentage is 9 points below, 84.4% and for upper secondary education the completion rate falls to 64.6% (UNESCO Institute for Statistics, 2023)

We employ the Overton tool to explore the evidence regarding Early Warning Systems (EWS) being implemented in education within the public sector in Latin America. The evidence is scarce, primarily found in documents from intergovernmental organizations (IGOs) that broadly describe the regional situation. These documents outline how various countries are beginning to implement systems and indicators to predict and prevent educational dropout. However, documents on this topic sourced from governmental think tanks and other national sources are less frequent. This scarcity makes it challenging to find detailed information about these systems and their implementations in each specific context, beyond the general descriptions provided by IGOs

Latin America is a diverse region, and studies like the one presented here contribute to a better understanding of how to track research influence on policy in the Global South and how these processes are represented on platforms such as Overton. Therefore, in addition to providing a general overview of this topic in the region, we will also focus on specific countries where these systems are known to be implemented. This dual approach will help to highlight both the broader regional trends and the particularities of individual national contexts.

This case study is structured as follows: First, there is a brief introduction to the background on the topic and critical terminology related to the case study. Next, we present the analysis of the policy documents conducted on a global scale, with a specific focus on Latin America. Subsequently, the analysis moves on to the underpinning research cited in the policy documents. Based on the findings, we depict Overton's advantages and shortcomings, which were identified while answering our research questions. Finally, we summarize the key outputs of this case study in the conclusions.

2. Background: Early warning systems (EWS) in education policy in Latin America

Social scientists have extensively studied educational exclusion and dropout risk, focusing on two main theoretical approaches: endogenous and exogenous factors within the educational system (Román, 2013). The aim has been to understand the causes and complexities in different regions. The digitalisation of administrative and learning processes has enabled new possibilities in terms of available data to analyze correlation

and trends associated with dropout. In line with this, computational sciences and engineering have evolved towards algorithms that use historical data to predict outcomes and find patterns in social behaviors. Particularly in education, in the last two decades, **machine learning techniques have been applied with multiple objectives**, one of which is **predicting dropout through early warning systems** (Mduma, Kalegele, and Machuve, 2019).

Early warning systems and indicators in Latin America are tools that are emerging thanks to digitizing processes in education ecosystems. Intergovernmental organizations and local governments promote them as effective means to monitor students' trajectories and prevent them from dropping out. These systems clearly demonstrate the use of data and new technologies for analysis and prediction, aiding in decision-making processes in public policies.

Early warning systems have been broadly used in the United States for the past fifteen years. Evaluation and results of EWS applied show positive outcomes when they are used together with timely interventions that support those students identified as at-risk of dropping out (UNICEF, 2018). Additionally, there are cases where the systems did not contribute to the main objective of decreasing dropout rates, and some showed biased results in specific demographics (Feathers, 2023). What is certain is that there is no one-size-fits-all model that meets the needs of every case. Each implementation has its strengths, weaknesses, and opportunities for improvement in planning, implementation, and the use and analysis of data (National Forum on Education Statistics, 2018).

These developments are still in their early years of design and implementation in Latin America, making the study of the underlying policy processes highly relevant. This involves reviewing the available policy documents that describe the problems in each territory, outlining how the different EWS are planned to be implemented, and evaluating whether they are integrated into a broader ecosystem for protecting educational pathways (Arias Ortiz et al., 2021). Additionally, when AI and automation are considered cost-efficient options for addressing high-impact social problems, it is crucial to critically analyze the cited academic research that supports these EWS.

Predictive systems rely on historical data to make projections or establish alert indicators. **Latin America is a profoundly unequal region, and educational indicators reflect the marked differences in performance and trajectories' results for different income levels** (Román, 2013; Kattan and Székely, 2015). Furthermore, social and human science research shows the importance of considering multiple variables when approaching students' performances and trajectories. In this context, it becomes relevant to pay attention to the cited research background to see if there is an interdisciplinary approach

that considers not only the necessary technologies to develop these systems but also the social, economic, and historical structures surrounding them.

Research questions

1. What evidence can we track of the influence on the design/adoption/implementation of Early Dropout Prediction Systems in education?
    - 1.1. What is the geographical distribution of policy documents?
    - 1.2. What is the evidence for Latin America?
    - 1.3. What is the underpinning research about/what research is cited?
    - 1.4. Who are the prominent donors/funders of relevant research?
2. What can we learn from other regions in the Global South to inform current work in specific cases in Latin America?
3. What is the gender representation in research platform metrics (policy and cited research)? Is there any bias/imbalance in the representation of women?
4. What are the advantages and challenges of using online policy research metrics to assess the influence of policy research?

3. Methodological considerations

We used Overton to answer the research questions and explored the results of policy documents and the underpinning research. The search began with exploring and defining keyword combinations that yielded results closest to the research topic. To do this, we started with the terms described in the terminology box below, 'dropout' and 'Early warning systems' in English and Spanish to approach the specific topic. Considering that EWS are a recent addition to education policy in Latin America and the availability of policy documents remains limited, we then expanded the search to include 'learning analytics' as a broader term, which provides for EWS but also other different uses of educational data as processes preceding algorithms and indicators to generate early warning systems.

Since this case study focuses on tracking research influence in incorporating EWS into public policy, we began the search by examining policy documents and their characteristics (using the ‘Search policy document’ tab in Overton). From these documents, we explored the underpinning research using Overton (through the

Explore/See the scholarly research these documents cite' feature) and by selecting and exporting results to work with outside of Overton. Figure 1 shows the main components of a policy document, a cited scholarly article in Overton, and the functionalities we used to obtain the information for this case study.

Explore/See the scholarly research these documents cite

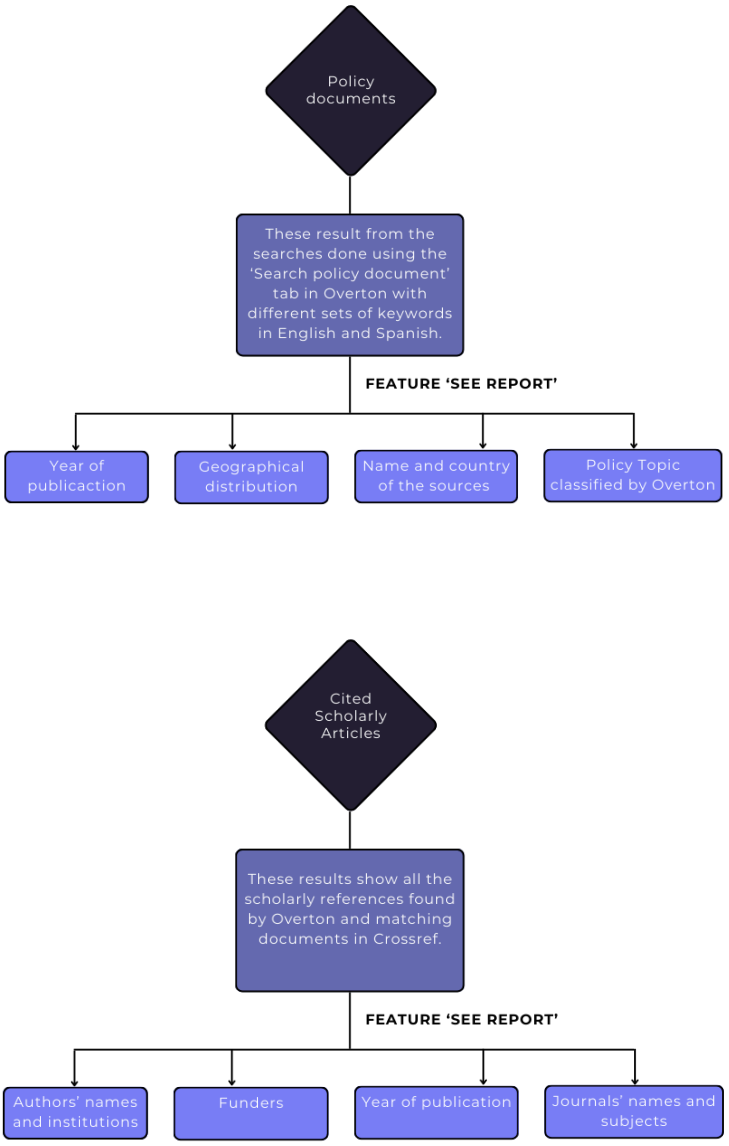


Figure 1: Description of main features of policy documents and cited scholarly articles results using Overton—source: authors' elaboration based on Overton's results and 'About the data' tab.

Terminology

Educational exclusion rate: this situation refers to those children or young people of school age whose family members responded in a household survey that they do not attend school (Acevedo et al., 2020).

Systems for the protection of educational pathways: Systems that aim at creating the conditions for continuous, completed, and high-quality educational trajectories for children and adolescents. Its two main components are detection and intervention (IDB, 2021). It is in the detection that EWS enters as a fundamental tool.

Dropout early warning systems (EWS): School dropout results from an ongoing process permeated by multiple factors. Some signals can help identify unstable/precarious schooling, which could result in a higher risk of dropping out. 'The EWS seeks to identify these warning signs early enough for schools and competent staff to implement the appropriate support that contributes to educational continuity' (UNESCO, 2022).

There are different types of EWS, and they can be separated into two groups: those based on expert knowledge or indicators and those based on machine learning algorithms (Arias Ortiz, 2021).

Educational management information systems (EMIS): These systems have the potential to collect, store, and analyze large volumes of administrative data, as well as data from students' learning processes. Data from EMIS and other sources can be used to develop EWS (either by finding trends and creating indicators or by training predictive models).

Learning Analytics is a field focused on collecting, measuring, analyzing, and reporting data about learners' processes and contexts. This data contributes to the optimisation of learning and the environment that surrounds it; methodologies within this field involve descriptive, diagnostic, and prescriptive analytics (SOLAR, 2011).

4. Analysis I: Research influence on early warning systems in education in the world

In this section, we use Overton to explore the evidence of influence on the design, adoption, and implementation of Early Dropout Prediction Systems in education. Specifically, we focus on the geographical distribution of policy documents.

The search using the combination **'early warning system*' AND education AND dropout'** without any filter brought back **1861 policy documents,** which included results mostly centered around education but also other topics such as health and agriculture among others. For this reason, we then applied the filter **'education'** in the filter 'policy topic', which resulted in **1160 results**.

Table 6: Geographical distribution of policy documents in Overton

Filters	Keywords: "early warning system*" AND education AND dropout
None	1861
Policy Topic: Education	1160
Policy Topic: Education Years: 2007-2024[2]	1061
Policy Topic: Education Years: 2007-2024 Source country: Central and South America and the rest of the Global South	76
Policy Topic: Education Years: 2007-2024 Source country: Central and South America	19 (**only four related to early warning systems**)
Policy Topic: Education Years: 2007-2024 Source: IGOs	504

Source: author's elaboration based on Overton's results.

[2] Overton's link to the search with those keywords and filters: https://app.overton.io/documents.php?query=%22early+warning+system%2A%22+AND+education+AND+dropout&year=_%3A97b&topics=Education&sort=relevance&added_before=2024-04-28

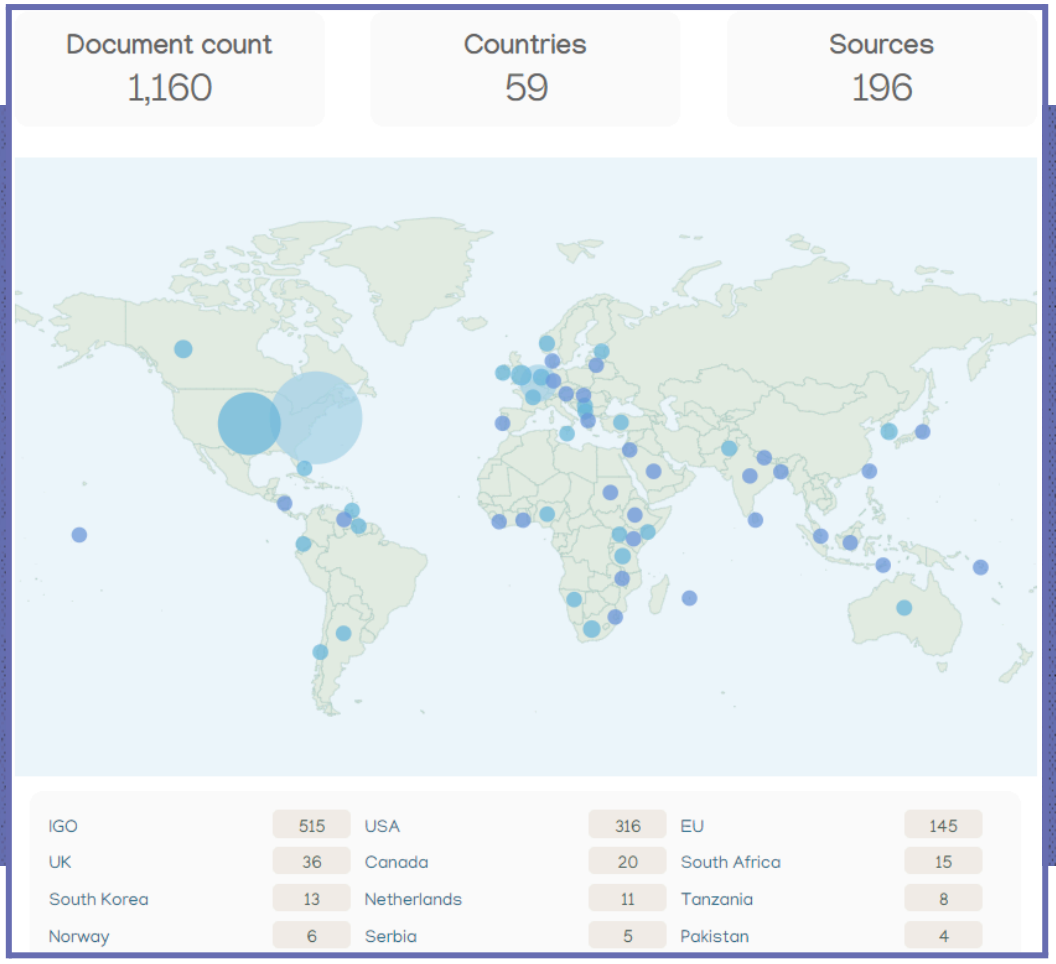


Figure 2: Geographical distribution of the policy documents matching the query '"early warning system*" AND education AND dropout' and with the topic Education (date: 19.03.24). Source: Feature 'See the report' in Overton.

4.1 Worldwide Analysis

The geographical distribution of the results shows that the **USA has 316 results,** which is the country with the **most policy documents.** The **EU** accounts for **145 papers,** and the UK, Canada, South Africa, South Korea, and The Netherlands follow, with 10 or more results each (Figure 2). **The results from the USA and the EU are the ones more closely related to school dropout and EWS or Learning Analytics to follow students' trajectories and prevent dropout.**

When considering results from other regions, **Uganda**, **South Africa**, **India**, and **South Korea** emerge as countries with a significant number of policy documents under these keywords. However, the majority of the papers from these countries are not directly related to Early Warning Systems (EWS) or technology applied to dropout studies in education. Instead, the policy documents are often related to diagnoses and reports on education, as well as other topics like the environment. This is despite literature showing

that there are cases where machine learning has been used to predict dropout in these regions (Bowers, 2021).

The results in the different regions of the Global South (excluding South and Central America) show education-related documents but not EWS. A few results in Asia relate Covid-19 to the use of technology in education. One policy document refers to a dropout prevention system applied in Timor Leste promoted by the U.S. Agency for International Development, which used data on behaviour, performance, and school attendance to identify students at risk of dropping out and provide support within the school.[3] Further research on tracking EWS in these regions could involve searches in local languages and checking for alternative terminology to refer to these processes in each territory.

Table 7: Policy documents results per region matching the search: '"early warning system*" AND education AND dropout', filtered by policy topic 'education' and relevant related topics.

Region (defined by Overton's filter 'Source region')	Results	Relevant associated issues of the policy documents
Africa	44	Education quality, Dropout rates, and Gender disparities
Asia	31	Dropout rates and prevention, COVID-19 responses, and EMIS.
Middle East	6	Dropout rates

4.2 Latin America

To track policy documents on EWS in Latin America, we use keywords in both English and Spanish. We considered documents sourced from governments, agencies, and think tanks in Latin American countries, as well as those sourced from IGOs (which, as pointed out by Overton, are geographically located by their headquarters, which are often located in the Global North).

In the search using English keywords ("early warning system*" AND education AND dropout) and filtering only policy documents sourced from Central and South American countries, we obtained 19 results, of which **only four were related to early warning systems**. Following this search, which focused on documents sourced from institutions within countries, we expanded our scope to include IGOs to see if there were any relevant

[3] School dropout prevention pilot project in Timor Leste.
Annual Report on Dropout prevention project from the USAID

results on this topic for countries in the region that may originate from organizations beyond local governments and institutions.[4] When filtering results sourced only from IGOs, we obtained 504 documents, most of which came from the **World Bank** and the **United Nations** and its various dependencies. Of those 504 documents, **25 were centered around Latin America and The Caribbean,** and 12 included the topics related to the search: Dropout prediction, Early warning systems, Education information systems, and Protected educational pathways. Of the 12 policy documents, **five are from the World Bank, four are from the Inter-American Development Bank**, two are from UNESCO, one from UN-Cepal and one from the Organisation for Economic Co-operation and Development.

When using the keywords in Spanish, the results of policy documents come mainly from Spain and the American continent (1189 policy documents). When filtering only results from **Mexico, Central and South America, we are shown 164 policy documents (See figure 3)**.[5] Similarly to what happens in the search using keywords in English, the documents most related to the topic are sourced from IGOs (most of which are the Spanish version of the IGOs' policy documents already shown in the search using the keywords in English). However, there is a discrepancy in the number of documents in both languages sourced from IGOs. While the results in English for this category are 504, the results of policy documents sourced from IGOs in Spanish are only 224.

A related topic that emerged linked to education reports in the searches in English and Spanish (both for policy documents sourced from Latin American countries and IGOs addressing issues in Latin America) was **the impacts of the recent COVID-19 pandemic on education in Latin America.** Specifically, how remote learning during the Covid-19 pandemic influenced the need to enhance different aspects of learning analytics, such as following students' educational pathways.[6]

[4] Overton's 'source type' classifies sources of policy documents in governments, intergovernmental organisations (IGOs), think tanks and others. Additionally, in 'source country, state, territory', it shows documents sourced in countries, specific territories and IGOs. When we include or exclude IGOs in the search to see results in specific regions or countries we do this in the geographical filter, not the 'source type'.

[5] Overton's link to the search using the keywords in Spanish and the selected filters by topic and territory: https://app.overton.io/documents.php?query=%22sistema%2A+de+alerta+temprana%2A%22+AND+educaci%C3%B3n+AND+%28deserci%C3%B3n+OR+abandono%29&added_before=2024-04-19&topics=Education&sort=relevance&source_country=_:b3f

[6] Latin America is the second region with the longest period of school closures due to the pandemic, with an average of 33 weeks of schools being fully closed and 37 weeks of partial school closure (Huepe, Palma and Trucco, 2022).

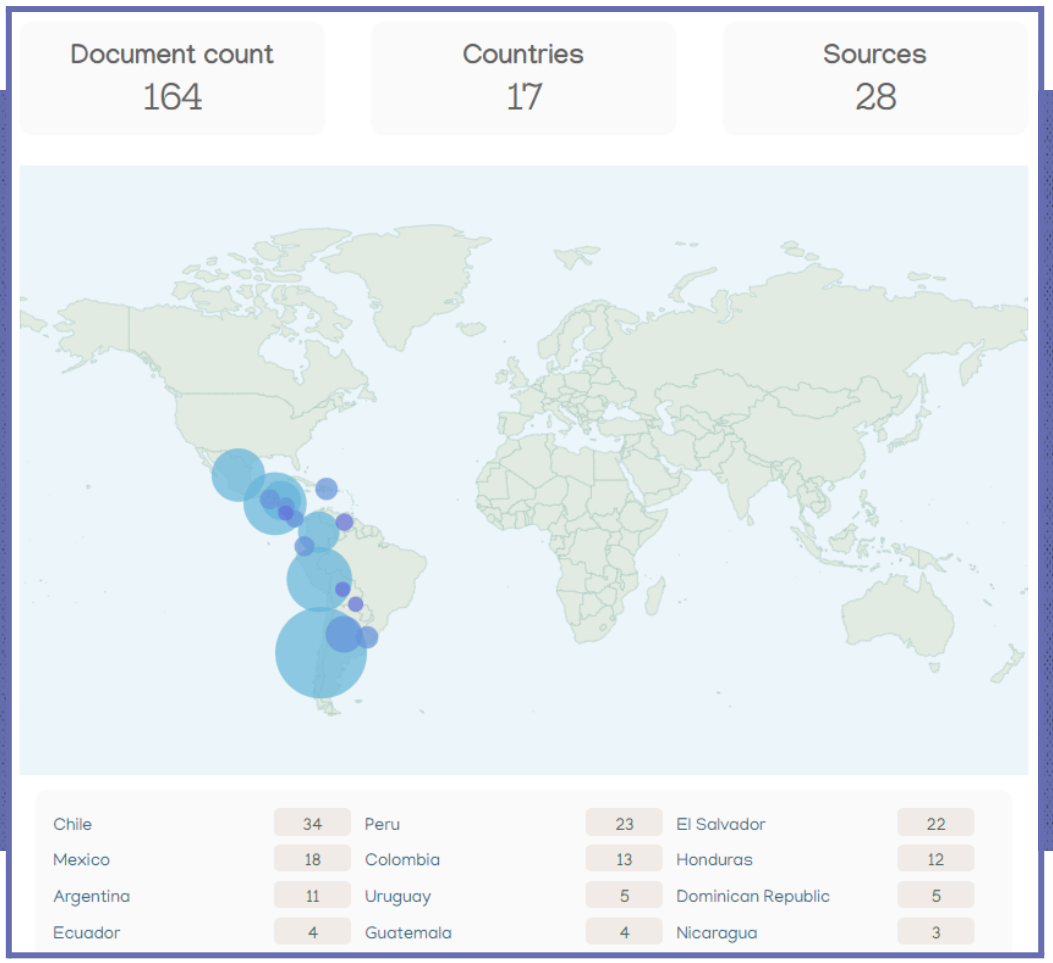


Figure 3: Geographical distribution of the policy documents matching the query '"sistema* de alerta temprana*" AND educación AND (deserción OR abandono)' and filtering the topic 'education' and 'source country, region, territory' countries from Latin America (date: 28.04.24). Source: Feature 'See the report' in Overton.

Overton is a valuable tool for uncovering international trends and specificities, for example, from Uruguay and Argentina.

Uruguay. The literature shows that EWS has been applied for some years in Latin American countries such as Uruguay (Arias Ortiz, 2021). In Overton, regarding policy documents sourced from governmental organizations, think tanks, and other entities in Uruguay, there are numerous documents covering topics such as digital learning, education data, learning analytics, and best practice guidelines. There are also National Protection System for Educational Trajectories documents, including indicators aimed at preventing dropouts from at-risk students. However, the **terminology in these documents does not include the term 'early warning systems'**. Therefore, although

they refer to early warning systems and indicators, they may not be displayed when the search is conducted using those specific terms as keywords

Nevertheless, the National Protection System for Educational Trajectories of Uruguay is mentioned and described in the policy documents sourced from IGOs on EWS in Latin America (reference is made to the Uruguayan case using that terminology despite not being used in the local context). **The search for policy documents from this country illustrates the need to contextualize the search. In other words, being aware of the differences in terminology around the same themes and issues in different territories is vital to approach the topics comprehensively.**[7]

Argentina. In Argentina, the implementation of EWS is still more recent; there are two provincial systems established in Mendoza and Entre Ríos. **This country is the only one for which policy documents on EWS applied in that territory (sourced from the government and IGOs) have been found in Overton with the given keywords.** However, we found policy documents for the EWS applied in Entre Ríos on governmental websites that were not found in Overton using the mentioned keywords.[8]

4.3 Broadening the search

Expanding our search beyond the specific terms "EWS" and "dropout," we broadened our scope to track **the geographical distribution** of documents related to similar topics under the umbrella of '**Learning Analytics**'. This encompasses not only Early Warning Systems (EWS) but also other processes that use educational data for learning and administrative purposes.

When we used the keywords: **'"Learning Analytics" AND predict* AND education' and filtered the topic 'education'** to see the geographical distribution of policy documents of related issues with a broader term, these were the results:

This search resulted in **667 policy documents**. Figure 4 shows that the **EU (25%) and USA (17%) are the territories with more documents in this search.** The publications shown for all available sources are from 2010 onwards (Figure 5). However, when filtering sources

[7] To date, Uruguay has 56,804 policy documents indexed in Overton. Szomszor and Adie (2022) place it as a country with a similar number of indexed policy documents as countries such as Brazil and Chile. However, the latter have a significantly higher production of scholarly articles.
Overton's link to policy documents sourced from Uruguay added before 3/05/2024: https://app.overton.io/documents.php?source_country=Uruguay&sort=relevance&added_before=2024-05-03

[8] Link to policy document found in government's website and not in Overton: https://cge.entrerios.gov.ar/wp-content/uploads/2023/03/SAT.pdf

only from countries in Latin America, there are 25 results from 6 countries, and the first year of the publications was 2014 (Figures 6 and 7). **These preliminary results align with arguments made in the academic literature.** These themes have been broadly studied in educational research, and their incorporation into public policies has been **present in the Global North** for more than **15 years**, while their implementation in the **Global South** is a **more recent phenomenon** of the last decade (Perusia and Cardini, 2021).

When looking at the results of this search, it is interesting to see the topics of the documents and how they relate to learning analytics, such as documents about higher education, online learning and massive open online courses (MOOCs), the use of AI and big data in education, ethical aspects and guidelines for the use of students' data, among others. Online learning and MOOCs serve as clear examples of the digitalisation of administrative processes and the rise of remote learning, which have facilitated extensive data collection and analysis for decision-making purposes. **Tools like Overton provide the capability to map related topics in policy documents, track the evolution of different themes and keywords over time, and compare various territories.**

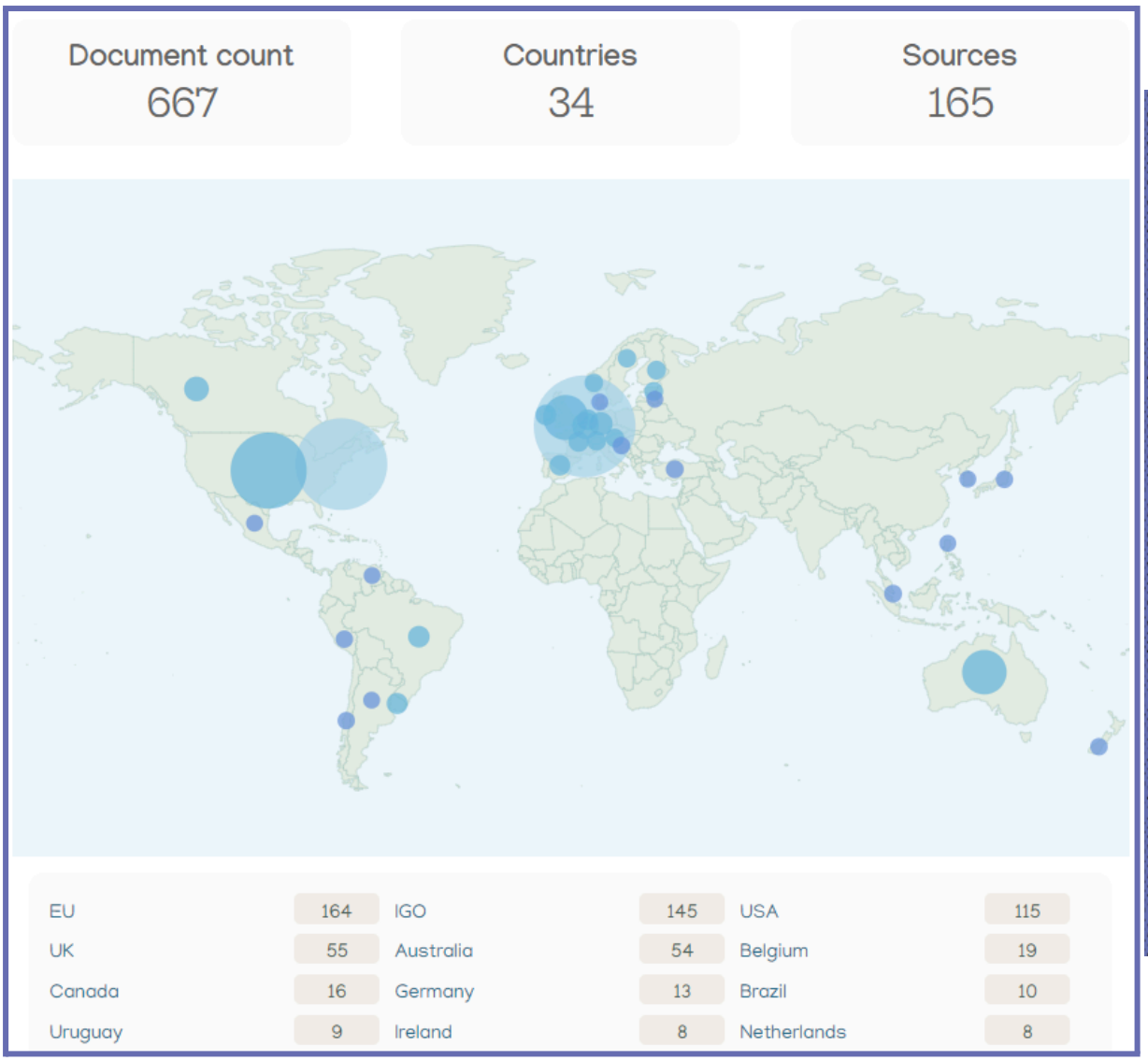
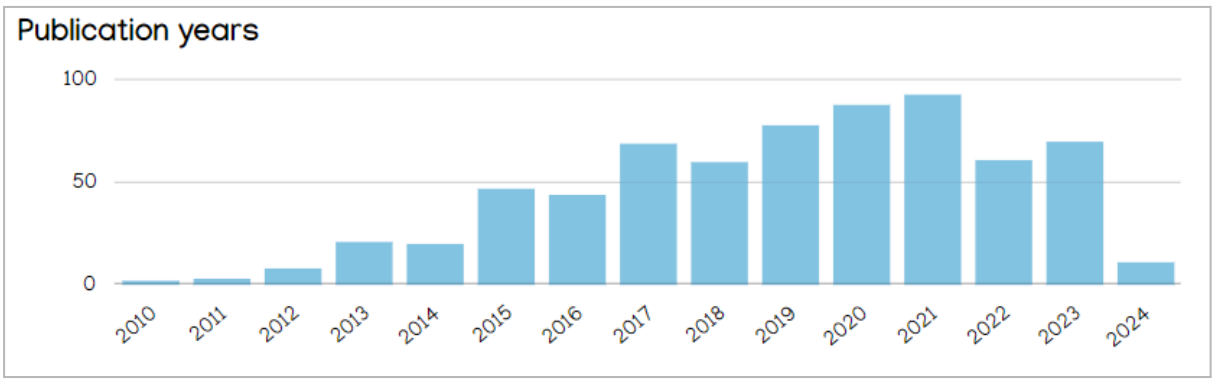


Figure 4: Geographical distribution of the policy documents matching the query '"Learning Analytics" AND predict* AND education' and filtering the topic 'education' (date: 28.04.24). Source: Feature 'See the report' in Overton.[9]



[9] Overton's link to the search of policy documents with those keywords: https://app.overton.io/documents.php?query=%22learning+analytics%22+AND+predict%2A+AND+education&topics=Education&sort=relevance&added_before=2024-04-28

Figure 5: Publication years of policy documents matching the query '"Learning Analytics" AND predict* AND education' and filtering the topic 'education'. Source: Feature 'See the report' in Overton.

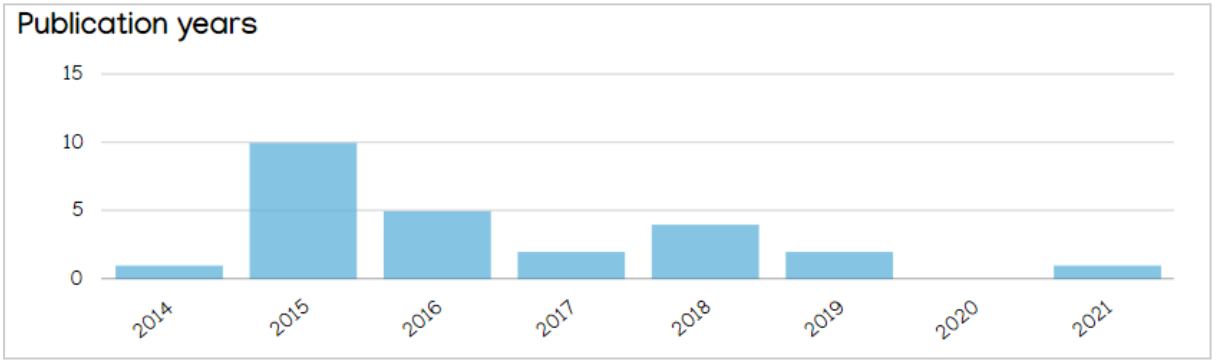


Figure 6: Publication years of policy documents matching the query '"Learning Analytics" AND predict* AND education' and filtering the topic 'education' and 'Source region' South and Central America (date: 28.04.24). Source: Feature 'See the report' in Overton.

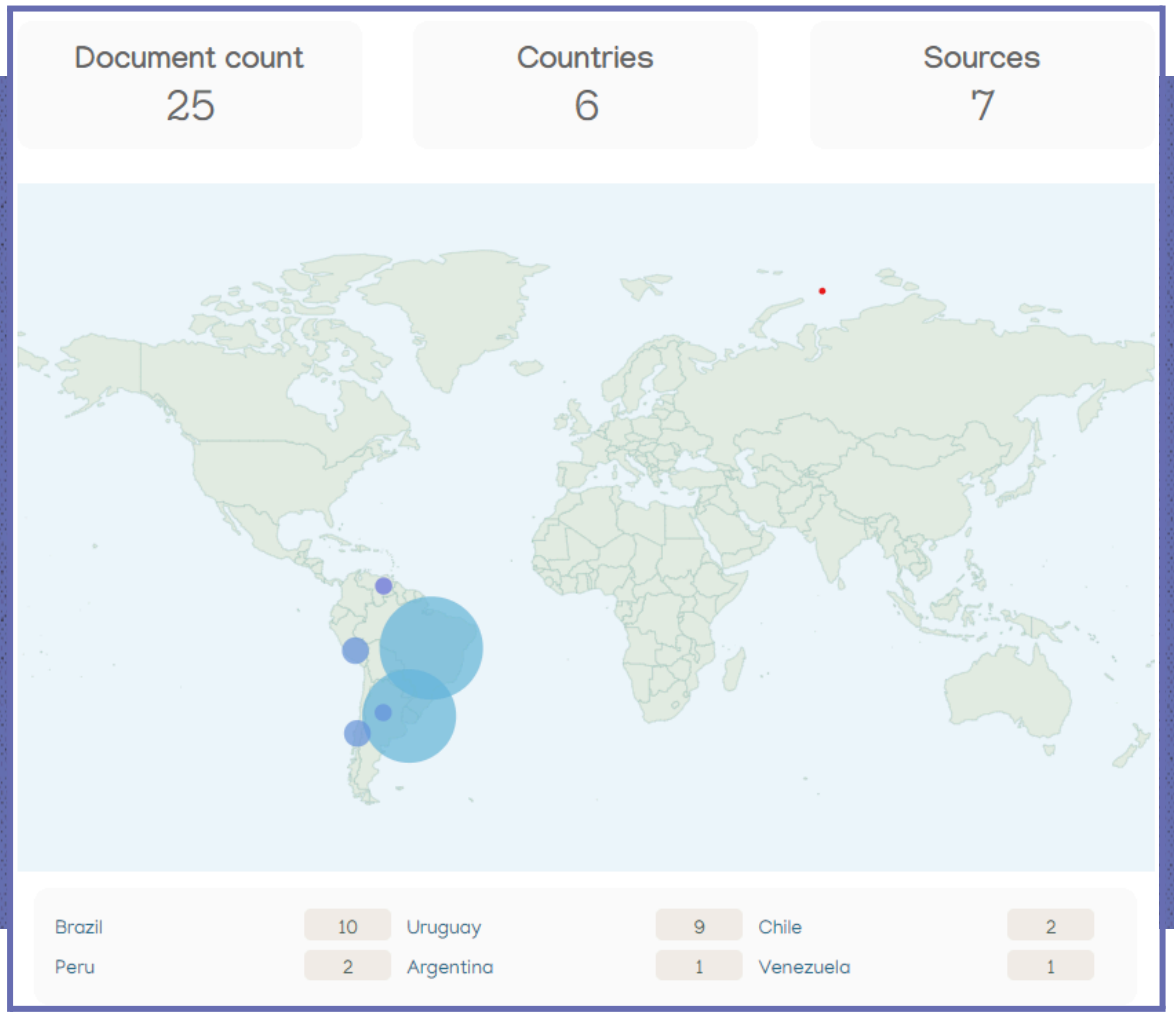


Figure 7: Geographical distribution of the policy documents matching the query '"Learning Analytics" AND predict* AND education' and filtering the topic 'education' and

'Source region' South and Central America (date: 28.04.24). Source: Feature 'See the report' in Overton.

5. Analysis II: Research influence on early warning systems in education in Latin America

5.1 Underpinning research cited by policy documents

Once we identified the policy documents about EWS, we investigated what underpinning research they cited[10]. First, we considered the 1061 policy documents from the worldwide search with the keywords and filters shown below (table 8). Our analysis revealed that these results covered a wide range of topics, including EWS, dropout prediction, and various education-related subjects. For this reason, following this general analysis, we conducted an underpinning research analysis of those policy documents resulting from searches sourced from governments, agencies, and think tanks in countries in Latin America, as well as those sourced from IGOs that explicitly address EWS to predict dropout in education in the region (sections 5.2 and 6).

[10] Overton's link to the scholarly articles cited by the 1061 policy documents: https://app.overton.io/articles.php?sort=relevance&cited_by_pdf_document_id=set:8026:f4ca781b8a570d78c66fe5e5ba94f984

Table 8: Keywords and filters used in the tab 'search policy documents, number of policy documents results, cited scholarly articles by those documents, main institutions and journals' subjects of the cited articles.

Keywords for policy documents search	Filters for policy document search	Policy documents results	Scholarly articles cited by those policy documents	Institutions with >200 cited articles	Journal Subject with >1000 articles
"early warning system*" AND education AND dropout	Policy Topic: Education Years: 2007-2024	1061	17991	World Bank Harvard Columbia Oxford OECD Stanford	- Education - Economic and Econometrics - Soc. and Pol. Science

- The years with the most cited research (more than 1000 articles) are 2018, 2019, and 2020.
- Regarding research **funders**, the **Economic and Social Research Council (ESRC), the National Institutes of Health (NIH), The US Department of Health and Human Services, and the National Institute of Health** funded more than 100 cited research documents. If we continue looking at the list of main funders, it is interesting to see that from **those with between 50 and 100 cited articles, 6 of the seven funders are medicine or health-related research funds**.
- When looking at the **authors' institutions of affiliation, the World Bank stands out with 658 articles cited, followed by Harvard University with 308 articles**. Other relevant institutions with more than 200 cited articles on these policy documents are Columbia University, the University of Oxford, the Organisation for Economic Co-Operation and Development (OECD), and Stanford University.
- When examining the presence of **institutions from the Global South among the first 100 institutions with 40 or more cited articles, we observe only five such institutions.** Among these, four are universities (University of Cape Town, University of Sao Paulo, University of the Witwatersrand, Aga Khan University), and one is an IGO, specifically the World Health Organization based in Pakistan.
- The **three categories of journal topics with more cited research in these policy documents come from journals on Education (1954), Economics and Econometrics (1748), and Sociology and Political Sciences (1247).** Several journal subjects with fewer cited articles can be grouped into Public Health or Medicine topics (6 of 9 subject areas that follow the first three subjects mentioned).

From the general overview, our next inquiry focused on investigating the evidence of research's influence on Early Warning Systems (EWS) in educational policies in Latin America (question 1.2). To carry out this action, we chose two sets of scholarly research cited by policy documents that allowed us to approach the set objectives with certain limitations. The **first set** of academic research studies was analyzed **using Overton's outputs, provided in the See Report feature**. The **second set** was analyzed **outside of Overton** because the policy documents used as references to observe the cited articles are not the result of a specific Overton query but rather a selection made retrospectively from a search yielding more general results. The following analysis will elucidate the decisions made for the study and present the results for each of the two sets.

5.2 Scholarly articles cited in policy documents from Latin America

These results are from the 225 scholarly articles cited by the four policy documents referring to EWS in the region resulting from the keywords '"early warning system*" AND education AND dropout', with the topic 'Education' and sourced from Latin America.

- Of the 225 results, **60% are classified as journal articles**, followed by books, which represent 16%.
- The years accumulating the highest number of cited documents are between **2020 and 2016**, with 73% of the total.
- The authors' institutions of affiliation with 4 or more articles cited are based in the **United States, Canada, and the United Kingdom**. Seven of them are universities, 1 is a company (Google), and 1 is an IGO (World Bank).
- There is only **one institution of affiliation from Latin America**: the **University of Sao Paulo**.
- The main funders of cited research for these policy documents are the program **Horizon 2020 by the EU** (with six articles) and the **Economic and Social Research Council** (ESRC) of the UK (4 articles).
- The main subjects of journals (with six or more results) in Overton's categories are the following:[11]

[11] The amount of cited articles about medicine and health innovations can be ascribed to one of the policy documents which analyzes different tools that apply AI in public healthcare systems around the world, as well as in other spheres in the public sector.

Table 9: Most cited journal subjects in the 225 results of scholarly articles cited by the four policy documents.

Journal Subject	Articles
Radiology, Nuclear Medicine and Imaging Economics, and Econometrics	13
Economics and Econometrics	12
Health Informatics	11
Health Policy	9
Multidisciplinary	9
Education	6
Public Health, Environmental and Occupational Health	6

The policy documents that cite the research documents analyzed in this section center around EWS in education policy in Latin America and the use of AI in the public sector in general in the region. However, **the cited authors' affiliations are predominantly from the Global North, except for the University of Sao Paulo, the only institution in Latin America**.

6. On Gender, location, and subject areas of scholarly articles cited in policy documents: using Overton outputs together with other tools

After exploring the underpinning research of the documents sourced from Latin America, this section analyzes the cited research that addresses EWS in education policy in Latin America in documents sourced from IGOs. This required exporting a group of 504 results from Overton (table 6) to perform later a customized selection, ultimately leading us to 12 policy documents with relevant content for this case. The underpinning research analyzed

in this section corresponds to the **237 research documents cited by these 12 policy documents**.[12]

- **According to Overton's classification, journal articles represent 59%** of the total cited documents. Books, with 14%, are the second category.
- The **years** with the most cited documents are **2020 and 2017**, accumulating 46% of the total cited research.
- **Stanford University** (14), the **World Bank** (13), and **Columbia University** (8) are the institutions with the most articles cited by policy documents.
- The **authors' academic profiles** share similar research interests and disciplinary backgrounds: Education **Economics, Education Policy, and Political Science**.
- Seven authors with more articles cited by these policy documents (3-4 articles cited), **six males**.
- The **countries** of the institutions of affiliations of these seven authors are the **USA, Germany and Canada**.[13]

Table 10: Authors of 3 or more articles cited by the selected policy documents. Source: Author's elaboration on most cited authors' information based on Overton's exports of cited scholarly articles.

Authors	Authorship of individual cited articles	Institutions	Country of Institutions	Gender	Disciplines/ Research interests
Ludger Woessmann	4	Ifo Institute for Economic Research	Germany	Male	Education Economics
Jason A. Grissom	4	Vanderbilt University, University of Missouri	USA	Male	Economy, Political Science, Education policy
Susanna Loeb	4	Stanford University,	USA	Female	Political Sciences, Economy, Education policy
Eric A. Hanushek	3	Stanford University, National Bureau of Economic Research, Ludwig-Maximilians -Universität München	USA, Germany	Male	Economy, Education policy

[12] List of cited scholarly documents from the 12 policy documents about EWS in Latin America: https://docs.google.com/spreadsheets/d/100ZWYYPtHOEFtcezO7_5i2tvEDL4ocX_/edit#gid=1711964642

[13] It is worth mentioning that 4 of the authors are associated by Overton with more than one institution.

Alex J Bowers	3	Columbia University, The University of Texas at San Antonio	USA	Male	Education Policy
John H Tyler	3	Brown University	USA	Male	Education, Economics, Public Policy
Philip Oreopoulos	3	University of Toronto, National Bureau of Economic Research, Canadian Institute for Advanced Research	Canada, USA	Male	Economics and Public Policy, Education policy

6.1. Subject Areas and Gender Analysis

This section focuses on the characteristics of the cited journal articles to delve deeper into the characterisation of subject areas and the gender of the authors. Since these scholarly articles were not obtained from a single search in Overton but rather from a selection made after exporting data from it, the results of this section were developed outside the platform, using other resources to characterize the cited research.

6.1.1. Subject Areas

Subject area analysis was carried out considering the categories used by Scopus to classify journals. The data of the cited research articles was merged with the Scopus Journal Classification Codes. In addition, gender analysis was carried out using Python's library "global-gender-predictor 0.0.3." After excluding those with missing data, we obtained 33 journal articles to analyze.

Scopus's 'subject area' categories are grouped into five 'supergroups': Life Sciences, Social Sciences, Physical Sciences, Health Sciences, and Multidisciplinary. Among the 33 scholarly articles analyzed here, the majority are published in journals categorized into Social Sciences (26), with 42% of this group falling under the subject area 'Economics and Econometrics', followed by 15% in Public Administration. There are two articles from Physical Sciences journals, one from General Computer Science and one from Artificial

Intelligence. There are also two articles with journals classified into both Social and Physical Sciences, one within the subject area 'Statistics, Probability and Uncertainty' and the other classified as 'Economics and Econometrics.'

Table 11: Subject areas and number of cited journal articles in each category. Source: Author's elaboration based on Overton's exports of cited scholarly articles and Scopus journals' classification.

Super group1	Super group2	Journal Articles	Subject Area	Journal Articles
Social Sciences	-	26	Economics and Econometrics	11
			Public Administration	4
			Developmental and Educational Psychology	2
			Sociology and Political Science	2
			Marketing	2
			Life-span and Life-course Studies	1
			Management of Technology and Innovation	1
			Law	1
			Organizational Behaviour and Human Resource Management	1
			General Business, Management and Accounting	1
Physical Sciences	-	2	General Computer Science	1
			Artificial Intelligence	1
Social Sciences	Physical Sciences	2	Statistics, Probability and Uncertainty	1
			Economics and Econometrics	1
Social Sciences	Health Sciences	1	Developmental and Educational Psychology	1
Social Sciences	Life Sciences	1	Life-span and Life-course Studies	1
Life Sciences	Health Sciences	1	Virology	1

6.1.2. Gender analysis

Throughout this case study, we have observed a limited amount of cited research from authors affiliated with institutions from the Global South. Even when the policies analyzed are implemented in these territories, most citations are to articles from institutions in the Global North. Additionally, the gender gap in scientific output is still a reality across subject areas (regarding professor positions, promotions, published papers, and citations). Studies suggest this disparity may be due to a productivity gap or a lack of acknowledgment of women's contributions (Ross et al., 2022). Therefore, it is relevant to include the gender variable in the research analysis cited by policy documents.

Out of the 33 journal articles used for the analysis, the "global-gender-predictor 0.0.3" tool inferred the gender of **100 authors**, identifying **36** of the names as **'female'**, **62** as **'male'**, and **two as 'unknown'.**[14] In addition, **out of the seven authors** with the **highest number of citations** across all cited articles of the 12 policy documents (237), only **one** is a **woman**. The sample of cited researchers selected needs to be bigger **for significant results**. However, the distribution of results regarding the gender variable resembles that of studies on the topic (UNESCO, 2019). Nevertheless, it is necessary to consider that the gender gap varies according to the territory and the disciplines (UNESCO, 2019; OECD, 2021).

7. Lessons learned: Metrics to track research influence on educational policy in Latin America.

Using Overton, an online platform designed to track available policy documents and the impact of research on policy in Latin America, provided a series of valuable insights. The platform allows for centralized searches by theme and offers various filters to tailor the results according to specific requirements. However, we also encountered some limitations, both within the platform itself and regarding the availability of documents on the subject. This section summarizes the main advantages and shortcomings observed. It is important to note that the aspects mentioned here result from using the platform for

[14] Accuracy y comparación con dataset general de Overton

this particular study, which presents specific characteristics regarding theme, region, and temporality.

7.1 Advantages

- Overton is a **valuable tool to find topics** related to the primary search. For this case, some related issues were the COVID-19 pandemic's education consequences, massive online open courses (MOOCs), and the new possibilities associated with data collection and analysis and EMIS in the region as a prior condition to allow the development of EWS.
- In the context of this case study's topic, which is still recent in the region, and considering that public policy processes require some time to accumulate documents for analysis and evaluation, Overton's web crawling **allows for an up-to-date database with the latest documents**.
- The data notes provided by Overton on possible biases or limitations of the platform for specific queries are valuable insights to consider when deciding whether it is helpful to add certain filters.

7.2 Shortcomings

- The search using **Spanish keywords** shows many documents whose titles do not match the titles of those in Overton. When we see the documents as search results, the text that Overton assigns as the document's title corresponds to another characteristic (e.g. the institution's name, country, or the type of document). This makes it necessary to open the document individually using the URL Overton provided to view the title and content and assess whether it matches the search topic.
- When exporting the results in Spanish to visualize and organize them in Excel, a large majority lack a snippet, making it difficult to search for keywords to identify mentioned and related topics. Upon opening the downloaded documents without a snippet, some indeed lack a summary or introduction, such as ministry resolutions. However, many others that do contain these sections are missing a snippet in the downloaded Overton spreadsheet.
- A few documents are classified as both 'policy documents' and 'scholarly articles'. These appear not only as results in a search for policy documents but also when examining the scholarly articles cited by the policy documents.

- The results of policy documents based on the keywords were **not specific enough** to provide only relevant results for the topic of the study. Even after selecting keywords and applying filters, more general issues continued to appear, such as documents about education policy or AI applied in policy that did not necessarily include EMIS or EWS. The lack of precision in the results prompted us to seek alternatives for selecting only the relevant documents. One approach was to analyze the underpinning research from Overton's general results and make a customized selection of individual policy documents, viewing and downloading each export of cited studies one by one.
- The exports provided by Overton for analyzing the cited research offer essential data on the scholarly articles. However, unlike the exports of policy documents, they do not include the abstracts of the articles or classifications by themes or disciplines other than the journal in which they were published. Additionally, information about research funders is not provided, and entries categorized as books or reports often lack authors' data. Consequently, it was impossible to conduct a gender analysis for those cases using the available exports.
- Considering the **underrepresentation of the Global South** in terms of policy documents and the significance of documents sourced from IGOs with offices in Latin America for this case study, it would be beneficial and accurate to differentiate IGO documents by country or, at the very least, by the region of the offices that publish them when this information is provided in the document. This approach would enable more precise searches and more accurately represent the geographical distribution of policy documents sourced from IGOs.

8. Conclusions

This study aimed to track evidence of EWS implementation in education policy in Latin America. We employ the Overton tool to explore the research questions. Both policy documents and cited scholarly articles and their main features were considered in the analysis. Our findings revealed that evidence of EWS implementation is scarce, with the majority of papers sourced from **Intergovernmental Organizations**. These papers broadly describe the regions' situation regarding implementing systems and indicators to predict and prevent education dropout across various countries while also outlining general guidelines for the effective implementation of EWS.

The geographical distribution of policy documents in Overton aligns with the literature on the subject, which explains the presence of these systems for nearly fifteen years in the United States and Europe, along with their emerging application in Latin American

countries (Perusia and Cardini, 2021; Bowers, 2021). Evidence of policy documents from Latin America on this topic is scarce, which makes it challenging to find detailed information about this type of system and its implementations in each context. Moreover, we found that for these searches, it was crucial to contextualize the topic for different territories. We noticed that terminology varied in various countries around the same theme. As a result, selecting different keywords was necessary to comprehensively address the topic.

The underpinning research of the different groups of **policy documents analyzed primarily** originates from the **Global North** in terms of authors' institutional affiliations, with the USA and the UK leading in the number of cited articles associated with their top-tier universities, along with some IGO, like the World Bank and companies such as Google. This trend persists even when considering only policy documents referring to EWS in Latin America. Institutions from the Global South, and specifically from Latin America, represent only a small minority of the cited research.

Regarding the subject area of the underpinning research, most of the cited articles were published in journals within the Social Sciences, notably in Economics and Econometrics.

The **gender analysis** reveals that only **one-third of the cited authors are women**, and among the seven authors with the most citations, only one is a woman.

From the searches on this topic, we conclude that Overton is a powerful tool for tracking policy documents and finding connections between topics. Although results were limited for this specific topic, it might be more useful for more general searches, such as tracking the use of AI in the public sector or the evolution of educational policies in Latin America.

From the searches conducted on this topic, we conclude that **Overton** is a **powerful tool for tracking policy documents** and **identifying connections between topics**. Although the results were limited for this specific topic, Overton may be more useful for more general searches, such as tracking the use of AI in the public sector or the evolution of educational policies in Latin America.

For future studies, it would be interesting to perform the inverse exercise of what was done in this study. Instead of starting from a specific topic to track it and seeing how it relates to the public policy ecosystem at a more general level, one could begin with a general search like the ones mentioned and then narrow down the analysis to specific themes found as relevant trends from those general results.

Additionally, considering the persistence of the gender gap in scientific outputs, it is pertinent for upcoming research to continue analyzing the distribution of authors in platforms and metrics that focus on research impact.

References

Acevedo, I., Castro, E., Flores, I., Perez Alfaro, M., Szekely , M., and Zoido, P. (2020). Los Costos Educativos de la Crisis Sanitaria en América Latina y el Caribe. IDB. (online) http://dx.doi.org/10.18235/0002838

Almeyda, G., Díaz, E., Elacqua, G., Fernández Coto, R., Hernández Cardozo, J. C., Dueñas Herrera, X., Margitic, J. F., Thailinger, A. and Zoido, P. (2022). What We Have Lost during the Pandemic. In: How to Reboot Education Post-pandemic Delivering on the Promise of a Better Future for Youth. IDB. (online) http://dx.doi.org/10.18235/0004241

Arias Ortiz, E., Giambruno, C., González Alarcón, N., Pérez Alfaro, M., Pombo, C., and Sánchez Ávalos, R. (2021). The road to educational inclusion: Four steps to develop systems to protect educational pathways. Designing early warning systems. IDB.

Busso, M., Bassi, M. and Muñoz, J. S. (2013) Is the Glass Half Empty or Half Full? School Enrollment, Graduation, and Dropout Rates in Latin America. IDB Working Paper No. IDB-WP-462. http://dx.doi.org/10.2139/ssrn.2367706

Bowers, A. (2021). Early warning systems and indicators of dropping out of upper secondary school: the emerging role of digital technologies. In: OECD Digital Education Outlook 2021: Pushing the Frontiers with Artificial Intelligence, Blockchain and Robots. OECD Publishing: Paris. https://doi.org/10.1787/c8e57e15-en.

Feathers, T. (2023). False Alarm: How Wisconsin Uses Race and Income to Label Students 'High Risk' – The Markup. (online) themarkup.org. Available at: https://themarkup.org/machine-learning/2023/04/27/false-alarm-how-wisconsin-uses-race-and-income-to-label-students-high-risk.

Huepe, M., Palma, A., & Trucco, D. (2022). Education during the pandemic: an opportunity to transform education systems in Latin America and the Caribbean. ECLAC. (online) Available at: https://hdl.handle.net/11362/48782

Kattan, R. B. and Székely, M. (2015). Analysing the Dynamics of School Dropout in Upper Secondary Education in Latin America: A Cohort Approach. World Bank Policy Research Working Paper No. 7223, Available at SSRN: https://ssrn.com/abstract=2585798

Mduma, N., Kalegele, K. and Machuve, D. (2019). A Survey of Machine Learning Approaches and Techniques for Student Dropout Prediction. Data Science Journal, [online] 18. doi: https://doi.org/10.5334/dsj-2019-014.

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.

Perusia, J. C., Cardini, A. (2021) Sistemas de alerta temprana en la educación secundaria Prevenir el abandono escolar en la era del COVID-19. (Early warning systems in secondary education: preventing dropout during Covid-19) CIPPEC #233.

Román, M. (2013). Factores Asociados al Abandono y la Deserción Escolar en América Latina: Una Mirada en Conjunto. (Factors Associated with School Dropout in Latin America: An Overview). Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación, 11, (2). pp. 34-59. https://doi.org/10.1038/s41586-022-04966-w

Szomszor, M., & Adie, E. (2022). Overton: A bibliometric database of policy document citations. *Quantitative Science Studies*, *3* (3), pp. 624–650. https://doi.org/10.1162/qss_a_00204.

UNESCO Institute for Statistics (2023) (online) Available at: http://sdg4-data.uis.unesco.org/.

UNESCO (2022). Early warning systems for school dropout prevention in Latin America and the Caribbean. REALC/UNESCO: Santiago. (online) Available at: https://unesdoc.unesco.org/ark:/48223/pf0000380354_eng

UNESCO (2019). Women in Science. (UNESCO Fact Sheet No. 55, 2019) (online) Available at .uis.unesco.org/sites/default/files/documents/fs55-women-in-science-2019-en.pdf

UNICEF (2018). Early warning systems for students at risk of dropping out. UNICEF Series on Education Participation and Dropout Prevention - Volume 2 (online) Available at: https://www.unicef.org/eca/sites/unicef.org.eca/files/2018-11/Early%20warning%20systems%20for%20students%20at%20risk%20of%20dropping%20out_0.pdf

Methodological Appendix

Research questions and the use of Overton

In this appendix, we present in detail the research questions and methodology used to achieve the case study's results.

1. What evidence can we track of the influence on the design/adoption/implementation of Early Dropout Prediction Systems in education?
    1.1. What is the geographical distribution of policy documents?
    1.2. What is the evidence for Latin America?

We retrieved documents from a selection of keywords used in the 'search in policy documents' tab to answer the first research question. After experimenting with different sets of keywords to use in the search engine, we settled on the combination **'"early warning system*" AND education AND dropout'** to obtain results specifically related to this topic. The search spanned from the year 2000 until the first documents related to dropout data and prevention emerged in 2007. Following the initial search in English, we replicated the process using the same keywords in Spanish: '"sistema* de alerta temprana" AND educación AND (deserción OR abandono)'

Table 1: Keywords and filters used in the 'Search policy documents' tab and the number of policy documents retrieved from each search.

Filters	Keywords: '"sistema* de alerta temprana*" AND educación AND (deserción OR abandono)'
None	1189
Policy Topic: Education	439
Policy Topic: Education Source country, territory: Mexico, Central and South America	164
Policy Topic: Education Source country, territory: IGOs	224

Once these initial two searches were conducted in English and Spanish, we visualized the repeated themes in the titles and contents of the documents that were on other topics related to education but also related to the **specific search about dropout prediction in education using early warning systems**, we conducted another, broader search using the following keywords: **'"Learning Analytics" AND predict* AND education' and filtering the topic 'education'.**

As mentioned earlier, early warning systems and indicators are relatively recent additions to education policy in Latin America. Given their novelty, the availability of policy documents on this topic remains limited. The selection of specific keywords was made to visualize the geographical distribution of documents on a broader topic: learning analytics, which encompasses EWS and other uses of educational data as processes preceding algorithms and indicators to generate early warning systems

1.3. What is the underpinning research about/what research is cited?
1.4. Who are the prominent donors/funders of relevant research?

To explore the underpinning research (questions 1.3 and 1.4), we used the following keywords and filters to get the policy documents' results. We used the feature 'Explore' - 'See the scholarly research these documents cite'. The feature 'See Report' shows the information about the journals' subject areas and funders.

Table 2: Keywords and filters used in the 'Search policy documents' tab, the number of policy documents retrieved, and the number of scholarly articles those policy documents cite.

Keywords for policy documents search	Filters for policy document search	Policy documents results	Scholarly articles cited by those policy documents
"early warning system*" AND education AND dropout	Policy Topic: Education Years: 2007-2024	1061	17991

From the general overview, we delve into investigating the evidence of research influence on early warning systems in education within the public sector in Latin America (question 1.2). To accomplish this, we selected two sets of scholarly research cited by policy documents, allowing us to approach the set objectives with certain limitations. The **first set** of scholarly research cited by policy documents was analyzed **using Overton's outputs in the See Report feature**. The **second set** was examined **outside of Overton** because the policy documents used to observe the cited articles were not the result of a

specific Overton query but rather a selection made retrospectively from an Overton search yielding more general results. The following analysis elaborates on the decisions made and the results obtained in each of the two sets.

After examining the underpinning research cited in the 1061 policy documents from the general search, we narrowed our focus to policy documents concerning EWS applied in Latin America, sourced from institutions within the region. This refined search yielded 19 policy documents, of which only 4 (sourced from Argentina and Venezuela) were directly related to the research theme[15]. Subsequently, we proceeded to analyze the studies cited in these four documents, totaling 225 scholarly articles

These are the keywords and filters that lead to this selection:[16]

Table 3: Keywords and filters used in the 'Search policy documents' tab to analyze documents sourced from Latin America, along with the number of policy documents retrieved and the number of scholarly articles cited by those policy documents.

Keywords for policy documents search	Filters for policy document search	Policy documents results	Scholarly articles cited by those policy documents
"early warning system*" AND education AND dropout	Policy Topic: Education Source region: Latin America Source country: Argentina and Venezuela	4	225

The policy documents found in Overton regarding EWS applied in Latin America and sourced from institutions within the region amounted to only 4. To broaden our analysis and delve into the research supporting these projects, we opted for an alternative dataset. Policy documents sourced from IGOs, which Overton locates primarily in the Global North, including entities such as the Inter-American Development Bank, the World Bank, UNESCO, and the OECD, tackle this theme in the region. These documents provide not only guidelines based on the specific context but also examples of how different countries

[15] Overton's link to the 4 policy documents mentioned (16/04/2024): https://app.overton.io/documents.php?query=%22early+warning+system%2A%22+AND+education+AND+dropout&source_country=_%3Abb9&source_region=_%3Ar5&topics=Education&sort=relevance&added_before=2024-04-16

[16] Overton's link to the scholarly articles from the search (26/03/2024): https://app.overton.io/articles.php?cited_by_pdf_document_id=set%3A8026%3A17126e70029721c17536b7d0917cdd4a&sort=relevance&published_before=2024-03-26

implement these systems. Hence, we chose to track the research cited by documents from IGOs.[17]

Since the results we obtained through Overton filters were not all about this topic, we took the policy documents matching results from the keywords '"early warning system*" AND education AND dropout' with the topic 'Education'. We sourced from IGOs (504 results), which were then extracted in Excel. After a closer look into the titles and snippets, we filtered **12 documents that refer to EWS in Latin America. We then extracted the cited research of those documents to analyze it**, which resulted in **237 cited academic documents**.

Table 4: Keywords and filters used in the 'Search policy documents' tab to analyse documents sourced from IGOs that refer to EWS in Latin America, the number of policy documents retrieved, and the number of scholarly articles those policy documents cite.

Keywords for policy documents search	Filters for policy document search	Policy documents results	PD on EWS in Latin America	Scholarly articles cited by those 12 policy documents
"early warning system*" AND education AND dropout	Policy Topic: Education Source country, territory: IGOs	504	12	237

To analyze the subject areas of the journals for the 237 academic documents cited, we considered the categories used by Scopus to classify journals. The data from the cited research articles was merged with the Scopus Journal Classification Codes. It is important to note that not all the academic documents cited are journal articles, and in some cases, this category is unknown. After merging the data extracted from Overton with the list from Scopus, we were able to analyze the subject areas of the journals and the gender of the authors from 33 cited journal articles.

[17] In Overton, the search for policy documents can be filtered by different attributes, including 'source country, state, territory.' Since the platform does not distinguish the territory when the source is an IGO, one of the categories of this filter is IGOs, which are located by their headquarters in the Global North even if they are produced by their offices in other regions. Once we tracked the policy documents sourced from IGOs and found them relevant for the analysis, we decided to include them and consider them as documents from the Global South despite the classification done by Overton differs.

2. What can we learn from other regions in the Global South to inform current work in specific cases in Latin America?

To answer research question 2, we used the same keywords in the 'Search policy documents' tab ("early warning system*" AND education AND dropout), adding the filter 'Source country, state, territory' and selecting countries with most policy documents from other regions in the Global South (excluding Latin America). We also used the filter 'Source region' and explored the results from Africa, Asia, and the Middle East separately.

Table 5: Keywords and filters used in the 'Search policy documents' tab to analyze documents sourced from countries in other regions of the Global South, excluding Latin America.

Keywords	"early warning system*" AND education AND dropout
Filter 'Policy topic'	Education
Filter 'Source region' Africa	Africa, North Africa, Sub-Saharan Africa
Filter 'Source region' Asia	South Asia, East Asia, Southeast Asia, Western Asia
Filter 'Source region' Middle East	'Middle East'

3. What is the gender representation in research platform metrics (policy and cited research)? Is there any bias/imbalance in the representation of women?

Overton's database of 'People' contains authors of scholarly articles cited in the policy documents found by the tool. Authors are linked to their institutions of affiliation, and gender information is not provided. To perform this analysis, we used the list of 237 scholarly articles cited in policy documents and utilized Python's library "global-gender-predictor 0.0.3".[18] This library infers the gender of the authors using their

[18] global-gender-predictor 0.0.3

first name and data from the World Gender Name Dictionary 2.0. The possible predictions include male, female, or unknown (unisex or not found in data).[19]

4. What are the advantages and challenges of using online policy research metrics to assess the influence of policy research?

Throughout the entire Overton exploration process, we identified strengths and limitations based on the research questions to fulfill our objectives. It is important to note that the aspects mentioned in this section pertain to the experience of using the platform for these specific searches, which had certain characteristics related to the theme, region, and timeline.

[19] As a limitation, there are no other gender alternatives (such as transgender, intersex, or others) in the classification of the predictor. Along with the gender analysis, we also conducted the analysis of subject areas. After merging the data exported from Overton with that of Scopus, we were able to visualize the assigned gender of 100 authors from 33 academic articles.