IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 17 April 2024, accepted 27 June 2024, date of publication 3 July 2024, date of current version 30 July 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3422822

RESEARCH ARTICLE

Data-Driven Strategies for Achieving School Equity: Insights From Brazil and Policy Recommendations

EMANUEL MARQUES QUEIROGA[1,2], ELTON SARMANHO SIQUEIRA[2,3], CARLOS DOS SANTOS PORTELA[2,3], THIAGO DAMASCENO CORDEIRO[2], IG IBERT BITTENCOURT[2,4], SEIJI ISOTANI[2,4,5], (Senior Member, IEEE), RAFAEL FERREIRA MELLO[2,6,7], ROBERTO MUÑOZ[8], (Member, IEEE), AND CRISTIAN CECHINEL[2,7,9]

[1]Instituto Federal de Educação, Ciência e Tecnologia Sul-Rio-Grandense (IFSul), Pelotas 96.015-360, Brazil
[2]Center of Excellence for Social Technologies (NEES), Universidade Federal de Alagoas, Maceió 57072, Brazil
[3]Faculdade de Sistemas de Informação, Universidade Federal do Pará, Cametá 68400, Brazil
[4]Harvard Graduate School of Education, Cambridge, MA 02138, USA
[5]Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo (USP), São Carlos 05508, Brazil
[6]Departamento de Computação, Universidade Federal Rural de Pernambuco (UFRPE), Recife 52171, Brazil
[7]Centro de Estudos e Sistemas Avançados do Recife (CESAR School), Recife 50030, Brazil
[8]Escuela de Ingeniería Informática, Universidad de Valparaíso, Valparaíso 2362735, Chile
[9]Centro de Ciências, Tecnologias e Saúde, Universidade Federal de Santa Catarina, Araranguá 88040, Brazil

Corresponding author: Cristian Cechinel (cristian.cechinel@ufsc.br)

This work was supported by the Ministry of Education (MEC), Brazil, under Grant TED10974.

ABSTRACT Education plays a crucial role in shaping societies and driving progress for future generations. Ensuring an education system that offers equal opportunities to all students is essential to promote equitable education. However, many students face disparities that impede their access to fair educational opportunities. This paper delves into the significance of using data-driven strategies to achieve a more equitable education system. We examine the barriers and inequalities hindering students' access to high-quality education and explore how data can be instrumental in addressing these challenges. To this end, we leveraged open educational data from Brazil, identifying eight school-related factors associated with student performance. We present various visualizations and insights on how these factors can be employed to promote equity among schools in the country. Furthermore, we recommend the implementation of two public education policies aimed at improving digital governance in Brazilian education. In summary, this paper underscores the crucial role of data-driven approaches in advancing educational equity and offers practical guidance for researchers and policymakers alike.

INDEX TERMS Data-driven decision, digital governance, education, public policies, equity, equality.

I. INTRODUCTION

Education plays a pivotal role in shaping societies and fostering progress for future generations. However, numerous students face barriers and inequalities that hinder their access to equitable education [1], [2]. Ensuring an education system that offers equal opportunities to all students is essential to promote equality [1], [2], [3]. Therefore, adopting data-driven approaches to education is crucial to achieving a more equitable system in which all students have equal access to high-quality educational opportunities [4].

The notion of equity in education has evolved over time, but the current theoretical framework outlines how equity should be integrated into the educational system to achieve equal opportunities and access regardless of gender or sociocultural background [5]. In some sense, this also has to do with eliminating biases in the education system based on race, sex, sexual orientation, religion, or socioeconomic

The associate editor coordinating the review of this manuscript and approving it for publication was Orazio Gambino.

© 2024 The Authors. This work is licensed under a Creative Commons Attribution 4.0 License.
For more information, see https://creativecommons.org/licenses/by/4.0/

background from the educational process, as mentioned by [1], [6]. While equity-based education is widely discussed, there remains an underexplored aspect concerning the quality of the educational system's structural perspective within the public school network [7].

The achievement of school equality can be approached at various levels of granularity. This work proposes an intermediate level, which measures the educational context in which students are immersed by contrasting the characteristics of different schools within the same system and assessing the students' collective performance as a measure of school effectiveness [2], [7]. Our understanding of equity in the present work aligns with the concepts of Minimum Standards, Impartiality, Equality of Conditions, and Redistribution explained by UNESCO [8].

Student performance is largely considered one of the most important factors to measure the success of educational systems around the globe. Countries and organizations normally recur to national exams to assess students' knowledge in important topics (reading, math, science literacy) in order to establish baseline guidelines, develop educational policies, and/or improve learning outcomes. Examples of existing instruments and tests are the Programme for International Student Assessment (PISA) conducted by the Organisation for Economic Co-operation and Development (OECD) [9] and the SAT (Scholastic Assessment Test) used in the United States for college admissions [10]. Student performance is also considered to be one of the key aspects associated with dropout rates in schools [11] and is also affected by a number of different factors about students themselves (socioeconomic status, health, intelligence, motivation) and their school (classroom environment, instructional quality) [12].

In Brazil, a number of educational exams are conducted by the National Institute of Educational Studies and Research (INEP).[1] Examples are the Brazilian National High School Examination (ENEM) [13], which is a standardized test to assess knowledge and skills of high school students that is largely used for students get access to the public universities, and the National System of Basic Education Assessment (SAEB), that biannually assesses knowledge of students from 3rd grade of elementary school and to students in the 3rd year of high school in subjects such as math and Portuguese language. The results of SAEB are used as a benchmark to make informed decisions about the national education system. Brazil also collects data on various aspects of the education system (number of schools and teachers, enrolment of students, school infrastructure) through its National Census of Basic Education (CENSO).

The present work uses educational data collected from INEP in order to unveil factors that are strongly associated with student performance and can help to achieve educational equity in the context of schools. The discovered factors are then used to provide graphical information and reports to the schools and policy-makers regarding the status of a giving school in relation to others. Such information can then be used to detect weaknesses and trigger actions focused on improving school conditions in order to achieve equity in terms of minimum standards and equality of conditions [8]. Through the analysis of these datasets, our aim is to gain a comprehensive understanding of the educational system, identify areas in need of improvement, and enable targeted interventions. Additionally, we seek to establish a data-driven approach to inform administrative decisions, aiding in resource allocation with the goal of achieving greater educational equality according to the principles of impartiality and redistribution [8].

The remainder of this paper is organized as follows. Section II covers the related literature, and Section III presents the methodology followed to extract knowledge from the data. Section IV describes the results obtained from the factor analysis experiment. Section V discusses the main school factors associated with students' performance and how they relate to the existing literature. Moreover, this section also offers insights into how these findings can be used to help school managers and policy-makers make informed decisions based on evidence. Finally, Section VI proposes policy recommendations that would allow long-term implementation of educational policies based on evidence, and Section VII concludes the article.

II. RELATED WORK

High school education plays a crucial role in the development of young individuals, consolidating and deepening the knowledge acquired during childhood and primary education. Additionally, it stimulates the development of critical, creative, and independent thinking, preparing students for academic studies and their integration into the job market [14], [15]

However, it is concerning to observe the alarming rates of dropout and retention in high school, especially in developing countries [15], [16]. Students who repeat a school year are 3 to 7 times more likely to drop out, which is a significant factor contributing to school abandonment [17]. This reality demands effective measures to address the issue and ensure that young individuals successfully complete their education.

Furthermore, the education system plays a crucial role in the development of any society, but issues related to the process, such as dropout rates, failure, retention, and lack of equity, pose significant challenges. These challenges compromise the academic, personal, and professional development of students and also result in economic and social damages [18]. Below, we highlight five key points that demonstrate the importance of addressing these issues according to UNESCO [19] and the World Bank [20].

- Academic detriment: Dropout or school abandonment can result in significant gaps in students' learning, compromising their academic formation and hindering progress in higher education.
- Professional detriment: Completion of high school is essential for entering the job market, and dropout or

[1]INEP is a government agency responsible for developing educational policies in Brazil.

abandonment can severely harm students' employment prospects, limiting their career options and salary potential.
- Personal detriment: Education is crucial for individuals' personal and social development, and its absence can lead to social exclusion and issues with self-esteem, self-confidence, and self-fulfillment.
- Increase in school dropout: Dropout or abandonment can create a vicious cycle of school dropout, where students feel demotivated to continue studying and leave school prematurely, thereby increasing dropout rates.
- Economic and social harm: Inadequate education acts as an obstacle to a country's economic and social development by impeding the formation of a qualified workforce and limiting the potential for innovation and progress.

Therefore, it is essential to seek methods to reduce issues related to students' educational journeys in order to improve the quality of education, enhance the country's productivity and competitiveness, and ensure a more prosperous future for future generations [15]. In this regard, it is crucial for high school education to be viewed as a pivotal stage in individuals' development, and effective policies and practices should be adopted to address issues related to the equity of the educational system as a whole [16], [21].

In this sense, the use of data as a means to assist and underpin the decision-making process in education emerges as an alternative. This process employs quantitative and qualitative analysis to gather information and evidence, which are then utilized to formulate strategies and support decision-making by education administrators [22], [23].

Furthermore, the use of data analysis techniques in the educational context can assist not only in the decision-making process but also in the formulation of effective educational policies by identifying patterns, trends, and gaps in student performance [24], [25]. This process aims to help identify factors influencing students' academic performance, guiding policy development [26] and revealing correlations between academic performance and factors such as extracurricular activities and school attendance [27].

In this way, data-driven public policies are increasingly valued for evidence-based decision-making, providing reliable and up-to-date evidence [22], [23]. At the same time, applying data analysis techniques to education-focused policies helps ensure successful interventions and the desired impacts on student learning [28].

However, the use of data in education presents challenges and risks that must be addressed. Student data privacy must be protected to ensure its educational purpose and prevent the misuse of data for other purposes [29], [30]. Furthermore, ensuring the quality and reliability of data is crucial to support public education policies with accurate and up-to-date information [31], [32]. Additionally, data-driven education public policies need to address these challenges and risks to use data effectively and beneficially for education without becoming a new source of bias in the educational process [33].

With this in mind, many researchers have begun efforts to use national-level data for evidence-based education public policies [34]. In the context of Latin America, the implementation of evidence-based education policies is still in development, with challenges in using educational data to inform policies [35]. Furthermore, financial and technical resource constraints hinder the application of data mining techniques in education in many countries in the region [35].

In Brazil, the National Institute for Educational Studies and Research Anísio Teixeira (INEP) plays an essential role in collecting and analyzing educational data at the national level. INEP is responsible for monitoring the quality of education in the country and gathering information about student performance, school and teacher data, attendance rates, and dropout rates.

It's important to emphasize that data should not be seen as an end in itself but rather as a tool for identifying problems, developing solutions, and monitoring the results achieved. Data-driven educational policies should be complemented by a holistic approach that considers other important aspects, such as teacher training, school infrastructure, and the involvement of the school community.

By adopting an evidence-based approach and using data to inform decisions, strategies can be developed to promote quality education, reduce dropout and retention rates in secondary education, and provide better educational opportunities for students. The combination of data analysis, educational experience, and community involvement can contribute to the development of effective and impactful policies that aim to continuously improve the education system.

III. PROPOSED APPROACH

In this work, we aim to exclusively utilize open educational databases provided by the Brazilian government as a source for analysis and experiments. Figure 1 depicts the general overview of the approach.

As can be seen in the figure, the approach proposed here involves leveraging the open educational data offered by the Brazilian government to identify educational factors related to student performance. These factors can subsequently serve as a foundation for promoting educational equity among schools nationwide and offer guidance to decision-makers in shaping public policies.

In this regard, we focus on the databases of the National System of Basic Education Assessment and the CENSO. SAEB is conducted semi-annually to assess the knowledge of students from the 3rd grade of elementary school to the 3rd year of high school in subjects such as mathematics and the Portuguese language. Meanwhile, CENSO is the primary data collection on various aspects of the educational system, conducting a comprehensive search for data such as

FIGURE 1. Overview of the proposed approach.

the number of schools and teachers, student enrolment, and school infrastructure, among others.

The intersection of both databases aims to provide an exploratory perspective for identifying factors that can assist in decision-making, with a focus on educational infrastructure equity. This analysis seeks to offer a comprehensive view of the educational system, enabling the recognition of areas that need improvement and facilitating targeted interventions to enhance school conditions. Furthermore, by adopting an intermediate level of granularity, this study aims to capture the conditions of educational provision among different public schools.

The CENSO is one of the main components of the educational information system, carried out by the National Institute of Educational Studies and Research (INEP). It promotes the collection of statistical and educational data and information related to basic education in Brazil. Also, It covers all levels of education (Childhood Education, Elementary School, and High School) and modalities (Regular Education, Special Education, and Youth and Adult Education). Based on the information collected by questionnaires, INEP obtains data related to student registration and performance, including information on gender, classes, grades, and periods, as well as information on the physical conditions of school buildings and existing equipment, in addition to information on technical, administrative, and teaching personnel, by level of activity and degree of education. From the previous data, some policies have been established to correct regional imbalances and promote equity in the provision of public education, such as the National School Feeding Program (PNAE) and the Learning Acceleration Program.

From the 17 different tables of the CENSO database, we selected the school table that contains 237 attributes (infrastructure, number of toilets facilities, access to the internet, operating hours, among others) related to 224,222 basic schools in Brazil where approximately 27 million of students from Basic Education were enrolled.

Every two years, SAEB is implemented in all public and private schools in Brazil, located in urban and rural areas, evaluating the quality of basic education from the third year of elementary school to the third year of high school, prioritization of areas of knowledge of Portuguese (focus on reading) and mathematics (focus on problem-solving). In addition to the evaluation process, students answer a questionnaire about their socio-cultural characteristics and study habits. The SAEB has made it possible to compare students' performance considering grades and years [36].

The SAEB database contained 136 attributes from a total of 70,606 schools. The SAEB stores information about the student population and the schools as a whole. Examples of information available are average grades of the students of the school in the different subjects, school location (urban or rural, or indigenous territory), teacher training level, socioeconomic level of the school, whether the school is private or public, percentage of students of the school which participated in the exams, the size of the school, how often the school offers training courses to the teachers, teachers experience, among others.

In the SAEB database, the exams (in Math and Portuguese Language) are administered to students from the 3rd and 9th grades of elementary school and to students in the 3rd grade of secondary education. The grades achieved by the students varied from 0 to 500 and were categorized in

levels that represent their proficiency in each subject. The SAEB database contained the average grades the students of each school achieved on the different exams (there was no information related to each student, only the compiled information about the whole school).

Data from SAEB and CENSO were intersected using the school identifier and preprocessed to remove noise (absent or duplicated columns). The Final database contained 202 attributes about the schools. Table 1 presents the summary of all datasets used in the experiments.

TABLE 1. Database summarising.

Dataset	Attributes	Schools	Students
SAEB	136	70,606	Not applicable
CENSO	237	224,222	27 million
Final	201	62,688	Not applicable

IV. RESULTS

Factor Analysis (FA) is a multivariate statistical technique often used in Data Mining to reduce the dimensionality of complex data sets and identify relationships and patterns between variables [37]. FA is an exploratory technique that seeks to extract the most relevant information from a data set, reducing the number of variables that need to be analyzed and providing an overview of the dataset [38]. The use of FA in Data Mining is important because it can assist in generating knowledge and making decisions in various areas, such as finance, marketing, social sciences, medicine, and education.

The application of FA in educational data has allowed the identification of key factors that affect student academic performance, allowing the development of more accurate and efficient prediction models [39]. This technique has great potential to improve education quality and develop public policies aimed at improving academic performance.

In the present study, after data pre-processing, factor analysis was performed to extract the most important factors related to school conditions. Factor analysis is suitable for this kind of experiment, as it explores the relationship between a set of variables and identifies correlation patterns between them [37]. The analysis determines whether the information can be summarized into a smaller set of factors, where each factor can be defined as a linear combination of the original variables. Factor analysis is mainly used to summarize and reduce data. The analyses were performed using Jupyter Notebook and Python libraries (Pandas, Numpy, and SKLearn).

In FA, sample adequacy should be assessed using two methods: the *Kaiser* − *Meyer* − *Olkin* (KMO) test and Bartlett's Test of Sphericity. The KMO is a statistical test that suggests the proportion of variance of the items that can be explained by a latent variable. This index indicates how appropriate the application of Exploratory Factor Analysis is for the data set. As a rule for interpretation of KMO indices, values less than 0.5 are considered unacceptable, values between 0.5 and 0.7 are considered mediocre, values between 0.7 and 0.8 are considered good, and values greater than 0.8 are considered excellent. On the other hand, Bartlett's test evaluates to what extent the covariance matrix is similar to an identity matrix; that is, they do not have correlations with each other. This test also evaluates the overall significance of all the correlations in a data matrix. The values of Bartlett's sphericity test with significance levels ($p < 0.05$) indicate that the matrix is factorizable, rejecting the null hypothesis that the data matrix is similar to an identity matrix.

For this experiment, our sample is adequate to perform the factor analysis, with $KMO = 0.94$ and Bartlett's test with $p < 0.001$, both indicating that the data are factorizable. The eigenvalue' is an excellent indicator for determining the number of factors [37]. In general, an eigenvalue higher than 1 will be considered as a selection criterion. Figure 2 shows the visual representation of factors' eigenvalues and the number of suitable factors. According to the figure, approximately 20 factors can significantly represent the data.

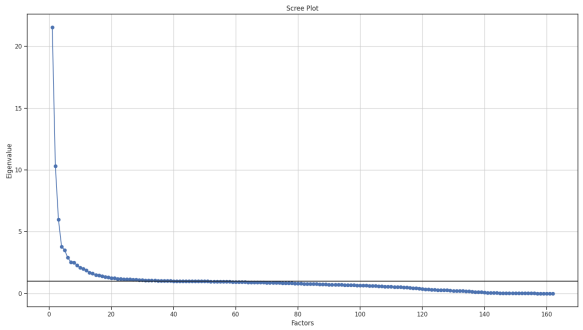

FIGURE 2. Sedimentation graph of eigenvalues after factor analysis.

Table 2 presents the total variance of the data that is explained by the factor analysis using the varimax rotation technique. As can be seen in Table 2, we have 35% cumulative variance explained by the 12 factors; this value shows the total amount of variation in a data set that is explained by these factors. Furthermore, SS Loadings can be used to determine how many factors should be retained in factor analysis. In our research, we retained factors whose SS Loadings were greater than 2, thus choosing the best factors that explain the variation in the variables. As noted by [40], determining the percentage of cumulative variance that adequately represents a dataset, along with setting thresholds higher than 1 for SS Loadings to select optimal factors, involves a subjective aspect of analysis and relies on contextual understanding. Moreover, in our factor analysis, we used Varimax rotation, an orthogonal technique that aims to minimize factor correlations, suggesting independence between factors.

In this case, we selected the 8 first factors from Table 2 to present here and use in the further implementation of services to the schools and policy-makers. These factors are School infrastructure (F1), Pedagogical Materials or Instruments (F2), Garbage Collection Treatment and Disposal (F3), LIBRAS and Braille (F4), People with Disability (F5),

Accessibility (F6), Internet access for the students (F7) and Students Transportation (F8). In the sequence, each one of these factors is briefly described:

A. SCHOOL INFRASTRUCTURE (F1)

The most significant factor associated with the performance of the students of a given school in the exams is the school infrastructure. This factor is associated with 25 attributes of the database, and it refers to the physical facilities, materials, and equipment that are required for the proper functioning of the school. Examples of the attributes composing this factor are drinkable water, public electric power system, public sewerage system, kitchen, computing laboratory, library, computer for students, covered outdoor space, and student food service, among others.

B. PEDAGOGICAL MATERIALS OR INSTRUMENTS (F2)

This factor deals with the pedagogical materials and instruments available at the school for the different classes. These materials may be of different types and for a variety of subjects, as well as focusing on different groups of students. This factor is associated with 11 attributes, such as multimedia material, child educational material, educational games, and sports materials.

C. GARBAGE COLLECTION TREATMENT AND DISPOSAL (F3)

Focused on the proper management of waste within the school premises, encompassing attributes related to garbage collection, treatment, and disposal. It includes binary indicators for the separation of recyclable waste, reuse of waste, recycling efforts, and the absence of waste treatment.

D. SIGN LANGUAGE (LIBRAS) AND BRAILLE (F4)

This factor deals with the support the school provides for students with visual and hearing impairments. LIBRAS stands for Brazilian Sign Language (Língua Brasileira de Sinais in Portuguese). Five attributes are associated with this factor, such as the existence of a LIBRAS interpreter in the school, videos with resources for hearing-impaired students, and Braille resources.

E. PEOPLE WITH DISABILITIES (F5)

Concerns encompass the presence of school infrastructure that adequately addresses the needs of students with disabilities, both physical and intellectual, as well as those with special educational needs, including gifted students. Key attributes constituting this include the availability of rooms and equipment tailored for individuals with disabilities, the presence of restroom facilities compliant with the Brazilian National Plan for Education (PNE), the number of accessible classrooms, and the provision of special classes specifically designed for students with disabilities.

F. ACCESSIBILITY (F6)

This factor represents the support the school is able to give to its students regarding accessibility. Three binary attributes (yes/no) are associated with this factor: the existence of ramps for wheelchair users, the existence of tactile flooring, and the nonexistence of accessibility at the school.

G. INTERNET ACCESS FOR THE STUDENTS (F7)

This factor describes if the school is providing internet access to the students. Two attributes are associated with this factor, which are the provision of the Internet on the school's computers and the provision of wireless connections to the students.

H. STUDENTS TRANSPORTATION (F8)

The last factor is associated with two attributes related to the provision of transportation so that students can arrive at the school. The two binary attributes (yes/no) are the existence of public transportation and the provision of buses.

V. DISCUSSION

Even though investments in education have increased in the last decade, Brazil still presents one of the highest unequal distribution of school funding among Latin American countries [41]. The recent literature has also indicated that variations in spending per student influence academic outcomes, with the positive effects being more prominent in disadvantaged student populations. Nevertheless, schools and regions characterized by lower income levels continue to experience persistent underfunding and a shortage of essential resources [41]. This suggests that school funding should consider regional disparities to ensure resources are allocated to those in greater need.

In addition to identifying disparities in the allocation of budgets and resources among schools, policymakers and educational managers should also assess the specific aspects where disadvantaged schools or regions require additional investments. In this scenario, the generated knowledge and revealed factors can serve as valuable tools to assist stakeholders in making informed decisions.

The majority of the factors identified here in the previous section align with existing literature in the field, showing their association with student performance to some extent. For instance, a study conducted by [42] revealed that the academic performance of primary education students in Latin America is affected by the presence of essential infrastructure (i.e., F1), including water, electricity, and sewage (i.e., F3). Besides, the presence of quality educational and didactic resources (i.e., F2), such as laboratories and libraries, along with the number of books and computers in schools (i.e., F2), was found to have an impact on students' achievement. The higher availability of computers per student in schools (F2) is also mentioned by [43] as associated with higher levels of student performance. Moreover, in the Turkish context, [44] found that higher quality of education resources in

TABLE 2. Total variance explained using varimax rotation.

	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12
SS Loadings	19.67	10.82	4.71	3.43	3.39	2.84	2.69	2.4	1.87	1.83	1.71	1.38
Proportion Variance	0.12	0.07	0.03	0.02	0.02	0.02	0.02	0.01	0.01	0.01	0.01	0.01
Cumulative Variance	0.12	0.19	0.22	0.24	0.24	0.28	0.29	0.31	0.32	0.33	0.34	0.35

the school (i.e., F2) are positively related to higher scores from students. It is crucial to emphasize that the associations outlined here are not universally applicable and appear to be constrained by specific contexts. As an illustration, [45] found that the condition of school facilities did not exhibit a direct relationship with student performance in the State of Wyoming, USA.

Similarly, the topic of student transportation (F8) is also addressed in the literature, yet drawing definitive conclusions regarding its impact on academic performance remains challenging and context-dependent. For instance, [46] observed that bus eligibility increased attendance rates among economically disadvantaged students in Michigan, USA, but no statistically significant evidence of effects on academic achievement was found. However, the author also acknowledged challenges in detecting such effects, particularly in relatively small groups of students. On the other hand, [47] noted that transportation challenges, including extended travel times, are linked to lower academic achievements, particularly for students attending low-quality schools. The authors underscored the significance of transportation in the context of student success, particularly among economically disadvantaged students. Besides, [48] identified correlations between students' distance to school and their academic performance in mathematics. The study recommended the formulation of public policies with a focus on infrastructure improvements to alleviate the challenges posed by long-distance travel for students. In this same direction, [49] points out that existing literature on school travel is weakly engaged with equity issues and that transportation equity is one of the requirements to promote educational opportunities and foster education equality. The authors propose that school districts and public agencies consider mobility in their educational policies.

The issue of Internet access for students in schools (F7) is perhaps one of the most controversial factors that emerged from our analysis. Various studies have reported divergent findings regarding the impact of Internet availability in schools. For example, in Malaysia, [50] discovered associations between internet broadband rates and the performance of both secondary and primary school students, with a stronger correlation observed in secondary schools. Conversely, in Florida, [51] suggested that the relationship between broadband subsidies and school performance could be considered negligible. In New Zealand, [52] noted that the implementation of fiber broadband led to an increase in success ratings among primary school students. Furthermore, the study provided evidence suggesting that schools with a greater proportion of lower socioeconomic students experienced more significant benefits from this enhancement. On the other hand, in Italy, [53] emphasized that access to high-speed internet connections in schools significantly decreased students' scores in Math and Italian language for the 8th grade. The authors specifically noted that students from low-educated parental backgrounds were the most adversely affected.

In the specific case of Brazil, in the study by [54], the authors utilized data from the SAEB in 2001 and concluded that, independently of socioeconomic classes, the use of computers and the Internet as pedagogical tools by teachers did not impact students' academic achievements. However, [55] later examined data from the "Prova Brasil," an expanded SAEB exam, and observed a shift in the effects of internet access at school between 2007 and 2009. The findings indicated a positive correlation between internet access at school and academic performance in 2009. The authors emphasized that while this influence was notable, the most significant impact was associated with internet access at home. They recommended that increasing internet access in schools should be accompanied by policy initiatives aimed at integrating internet-based educational tools into the curriculum. This is an important aspect of internet access in schools as occurrences of students misusing this resource to engage in social media and gaming activities rather than utilizing it to support learning have contributed to the controversy surrounding the topic [56].

Nonetheless, to the best of the authors' knowledge, this is the first occurrence where all these factors appear combined within the same article and arise from empirical data.

A. INFORMED DECISIONS BASED ON KNOWLEDGE

This section provides insights on how the generated knowledge can be used to help policymakers understand the overall educational scenario of the country and to make informed decisions toward equity based on evidence. The main idea is to use the factors that are associated with student achievement (and that reflect the current state of education in the nation) in order to provide visualizations at different levels of

granularity for policy-makers and education public managers. These visualizations will enable them to identify the primary areas requiring investments and actions at different levels (i.e., country, state, and city) and from the perspective of equality across schools.

1) COMPARING CONDITIONS OF THE SCHOOLS AMONG REGIONS

Once the important dimensions for good student performance in schools are calculated, the government and decision-makers are able to contrast the conditions of the schools in the country according to each criterion. For instance, in the search for equity, the national government may compare the conditions of the different regions of the county in accordance with the dimensions and establish priorities of attention for those regions that require investment in aspects that are being under-served. As it can be seen in figure 3, the regions present huge differences in two factors that can be considered crucial for equality which are School Infrastructure (F1) and Accessibility (F6). For the case of F1, it is clear that the Southeast region is the one with the best conditions, while all the others are far behind (especially the Central-West and North regions).

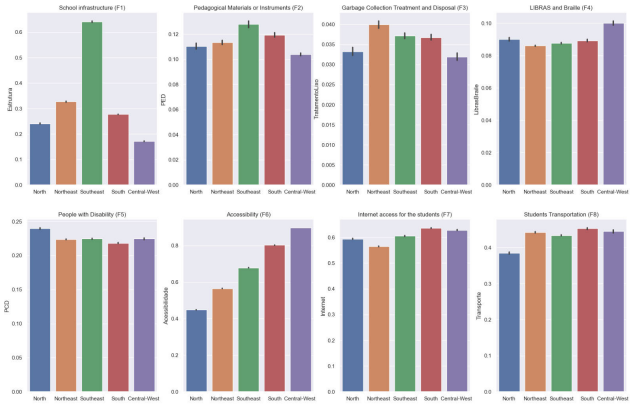

FIGURE 3. Comparison of the Brazilian regions in accordance to the factors.

It is important to highlight that this kind of comparison is also possible at different administrative levels, i.e., it is possible to perform comparisons among states (federative units) from a given region (or all states of the country), among cities of a given state, and schools of a given city.

Moreover, interactive maps can help enhance knowledge about the need for investments in specific geographical areas, considering a given factor of importance. Managers and policymakers are able to overview the most problematic regions to oversee their general conditions. Figure 4 reinforces the previous comments about how the Southeast region currently has the most favorable School Infrastructure conditions (F1).

2) ASSESSING IMPLEMENTED PUBLIC POLICIES

The knowledge can also be used to follow and assess the results of previously implemented public policies.

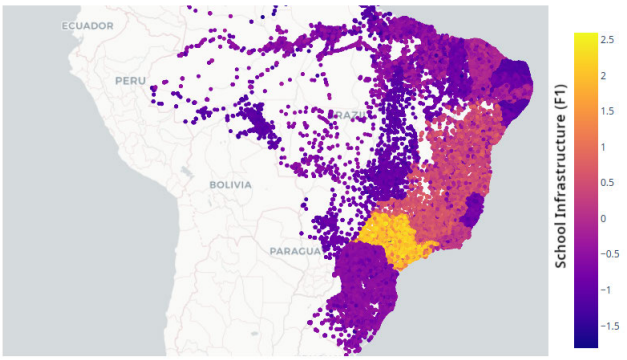

FIGURE 4. Countrywide overview of School Infrastructure (F1).

For instance, in 2008, the Brazilian government implemented an initiative called the Broadband in Schools Program (PBLE, from the Portuguese Programa Banda Larga nas Escolas). The PBLE intended to provide high-speed internet access to public schools across the country in order to create a more connected learning environment in the schools. As it is possible to see in figure 3, the Internet Access for the Students factor (F7) is very balanced among the regions, showing a high level of equity among the schools across the country. This can be confirmed if we depict this factor across the several states of the country (see Figure 5).

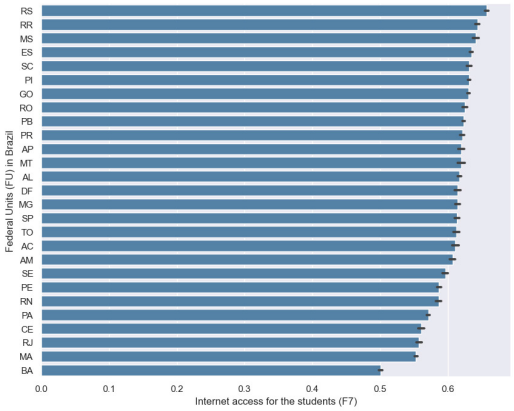

FIGURE 5. Comparison of the states in accordance to F7.

It is important to highlight that information in Figure 5 does not mean that all schools in the country are in the same conditions regarding access to the internet by the students. As the information provided here is an average of the factors inside each state, it is possible to have those schools that are way under this average and thus in need of investments in this factor. However, such visualization helps the manager compare the average conditions among states and have a broader picture of the factors in the whole country. Furthermore, there is also the opportunity to delve deeper into data visualization, allowing one to scrutinize the conditions of cities within states and schools within cities.

3) PINPOINTING SCHOOLS

The interactive geographical map can be used by policy-makers and managers to select specific schools and evaluate their conditions in accordance with the different factors. Figure 6 shows the selection of a given school (let's call it here school x) in the map and the presentation of different information about the school (location, urban or rural, federal unit, restrictions, level of education, among others). As can be seen in the figure, the board is colored according to the specific factor under analysis. For example, this school presents good infrastructure conditions (F1).

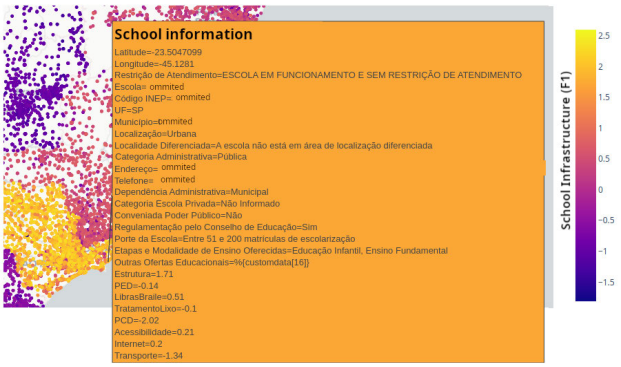

FIGURE 6. Pinpointing a specific school in the map.

Once the school is selected, it is possible to recommend investments in specific aspects of the school to improve its conditions in the factors that are underserved in comparison to other schools. In that sense, in a hypothetical situation where the manager has a certain budget for resource allocation in an individual school, city, or even a network, the system can provide various options to enhance the institution's quality and create a more suitable educational environment.

This is accomplished through a basic recommendation system that uses the average and standard deviation for each of the factors generated for the school in relation to the same factors generated for the state. Therefore, if a factor is identified as below the state average, it is recommended to prioritize investments in it. Figure 7 shows the recommendations for investment for a given school in Brazil. The distribution of the investment will depend on the specific needs of the school and the priorities established by the manager in conjunction with the pedagogical team. A strategic approach and careful planning are essential to maximize the benefits of the investment and promote significant improvements in the quality of education offered by the school.

In the Figure, factors that are under the state average are colored in red. For these factors, investment in specific aspects is suggested for the education manager. For example, for the school selected in the figure, the system recommends investing in five out of the eight factors. It is important to highlight that similar recommendations can also be applied on a larger scale, such as at the city, state, or regional level.

4) DELVING INTO SCHOOLS' FACTORS

More in-depth information can be provided to the managers so that it is possible for them to understand how far from the state average each one of the schools' factors is. Figure 8 presents a speedometer graph where the manager can observe the details of each factor of the school x. It is interesting to compare information here with information from the previous Figure 7 in order to understand how the manager can decide to invest more in one factor than in another or how to intelligently distribute investments among the different factors.

As it is possible to see, in Figure 7, School Infrastructure (F1) is considered to be below the average and received a recommendation of investments. This information is confirmed by Figure 8; however, it is possible to see how close to the boundaries of the state average this factor is. The same does not occur with Pedagogical Material or Instruments (F2), Garbage Collection Treatment and Disposal (F3) and Sign Language (LIBRAS) and Braille (F4), which are way below the threshold of the state averages. This information may help managers decide the amount and distribution of investments needed to increase the quality of each factor of the school in order to achieve the averages of the state.

VI. POLICY RECOMMENDATIONS

The rise of digital technologies and new sources of data have opened the possibility for governments to address policy issues using data-driven approaches. Innovative practices of data collection, sharing, and usage are emerging as a way to optimize decisions and the development of more solid and evidence-based education policies [57]. There is currently a growing trend towards the adoption of new digitized tools that provide real-time insights from the continuous flow of data. This is particularly important for education policy-makers who can be able to base their decisions on evidence that emerges from real educational settings [58]. This scenario will allow the emergence of the so-called anticipatory approaches for education governance (or anticipatory governance), where new forms of techniques (mostly consisting of Artificial Intelligence, Predictive Analytics, and Statistical methods) will help to shape the future of education and other sectors and domains by allowing to policy-makers to foresee trends and needs [59], [60].

Given these considerations, we emphasize two primary policy recommendations that would allow a long-term educational policy implementation based on data and evidence: 1) The establishment of a robust data quality infrastructure, and 2) The encouragement of initiatives focused on measuring and promoting equity in education achievement.

A. ESTABLISHMENT OF A ROBUST DATA QUALITY INFRASTRUCTURE

The establishment of a robust data quality infrastructure is demanding global attention and is highly associated with the rise of new services and modes of governance [61].

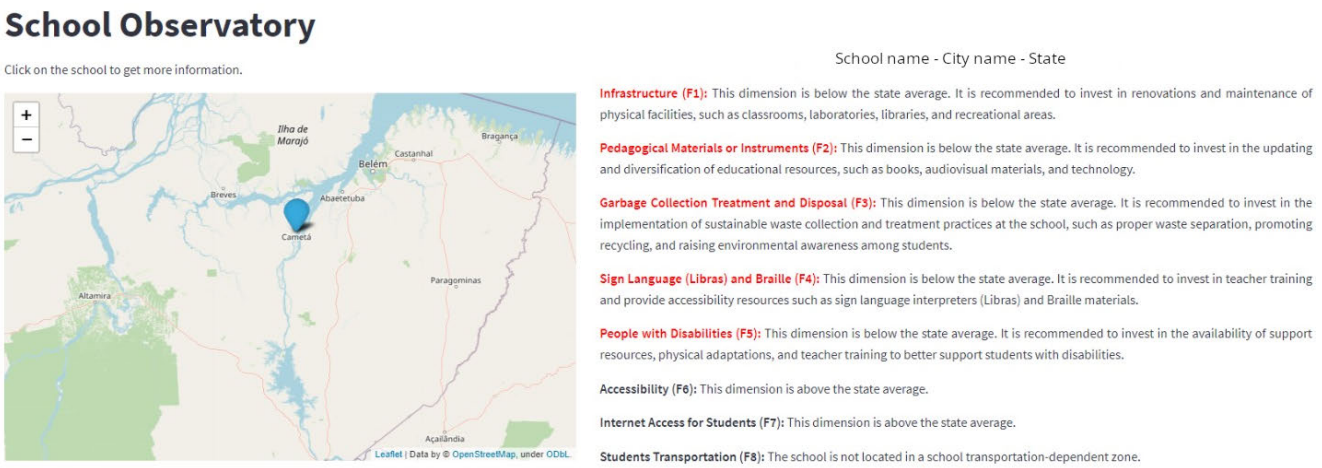

FIGURE 7. Recommendations of improvements in the school.

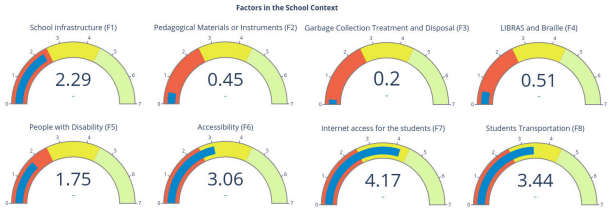

FIGURE 8. Speedometer graph showing each factor of the school in comparison to the state average.

According to [61], data infrastructure to support data-driven education services is mostly interrelated to three fronts which are: the *scales* across which the infrastructure is being integrated, the *scope* of the data that is being created and the *explanatory power* that is enabled by the data infrastructure. In order to foster data-driven public policies in education in Brazil, data infrastructure should be improved on these three mentioned fronts. In the case of Brazil, there is still no single personal database in the educational context, and each public organization is responsible for maintaining, securing, and managing its own data [62]. Building quality evidence and making sense of this evidence (policy quality systems) is pointed out in the literature as one of the capabilities for successful implementation of transformative policies (together with leadership, engagement, and people capability) [63]. Therefore, it is crucial to step further towards building policy capabilities grounded on the existing systems, data, and instruments currently available, as well as to evolve the existing legislation that allows such developments.

A poor data infrastructure leads to low-quality data that may distort the real conditions of education at all different levels, thus leading to inadequate decisions and policies. The existence of open educational data in Brazil, together with continuous and structured policies for gathering structured educational data, allows the creation of data-based tools and strategies to guide and improve public policies in this sector [62], [64]. However, efforts should be made to integrate and share data among different public organizations in Brazil. It is necessary to integrate the existing large-scale and longitudinal datasets, as well as develop data dashboards and services to consume the integrated data [65]. This could be partially done in a similar way as the Australian NISP initiative [66], [67], where Federal, State, and Territory Ministers for Education are together enforcing interoperability standards (and a Learning Services Architecture) to guarantee uniform data formats across all schools in the nation. In Brazil, such integration should comprise the different education and administrative levels, covering students' performance and historical data, students' socio-demographic data, data about school conditions (such as the ones tackled in the present paper), as well as large-scale assessments conducted by the national government. Lastly, to ensure the production of high-quality data, it is imperative to allocate resources to enhance the expertise of professionals specializing in education data management and data literacy. These experts should possess the skills required to efficiently gather, manipulate, and process data to derive meaningful insights and knowledge. The existing data infrastructure, once integrated, will allow a historical analysis of the students and schools and pave the way for the development of artificial intelligence-powered tools able to provide real-time policies. This new data infrastructure topology will help to emerge a proper data education ecosystem with increased explanatory power about the different education dimensions in the country. Such an ecosystem will also enable all kinds of interventions in education, including those focused on promoting equity.

B. ENCOURAGEMENT OF INITIATIVES FOCUSED ON MEASURING AND PROMOTING EQUITY IN EDUCATION ACHIEVEMENT

This study underscores the importance of equity-based approaches in education to promote a more inclusive and

fair educational system. By acknowledging and addressing the barriers and inequalities that hinder students' access to quality education, we can pave the way for a future with less unequal societies. Through an intermediate level of granularity (school level), we explore the educational context in which students are situated, examining school characteristics and their impact on overall performance. Our analysis, based on open educational data from Brazil, has shed light on factors related to educational equity and provided valuable insights for generating initial knowledge on the subject.

However, the most detailed level of granularity in the available data is at the school level. This limitation makes it impossible to verify how those factors (conditions of the school) are effectively reaching those students in need of assistance and to detect the required school policies for improving quality and fostering equity. For instance, it is possible to assess whether the school is effectively addressing the factor of accessibility, but it is challenging to quantify the extent to which students with disabilities are being adequately served. In order to verify that, it is essential to explore a finer level of data granularity, specifically at the student level. It is well known that students' performance is not only associated with school attributes but is also dependent on factors such as students' socioeconomic vulnerability, family support, students' affective, psychological, and cognitive aspects, and the relations of the school with the community [68]. Moreover, it is also expected that interventions focused on promoting equality in the whole school and considering different background aspects tend to have a positive impact on education quality [69]. Given this, it is crucial to provide schools with the necessary resources to identify the various factors influencing student performance within their specific context. These resources encompass more than just effective data collection; they also involve empowering stakeholders to engage in educational research that generates fresh insights, as outlined in [70]. This empowerment will facilitate the development and execution of specific improvement strategies and intervention plans intended to elevate the quality and equity in education. These plans should be firmly rooted in dependable knowledge and information, making them easier to monitor and assess.

VII. FINAL REMARKS

The present paper describes how educational data at the school level can be used to deliver evidence-based recommendations for educational managers. For that we integrated data from two distinct databases available in Brazil (CENSO and SAEB) and extracted 8 factors associated to students performance (School infrastructure (F1), Pedagogical Materials or Instruments (F2), Garbage Collection Treatment and Disposal (F3), LIBRAS and Braille (F4), People with Disability (F5), Accessibility (F6), Internet access for the students (F7) and Students Transportation (F8)). These factors were then used as baseline information to allow informed decisions based on knowledge. We showcased various examples of visualizations representing different administrative levels. These visualizations are intended to assist education public managers in identifying disparities among schools, states, and regions. They can use this information to address weaknesses and attend to specific needs that require attention. Together with the practical examples of possible use of the educational data, we also recommended the implementation of two education policies to improve government capacity in constructed data-driven decisions: 1) The establishment of a robust data quality infrastructure and 2) The encouragement of initiatives focused on measuring and promoting equity in education achievement.

It is known that education data is gainer more and more attention from governments that need to overcome old bureaucratic approaches and are required to take real-time actions grounded on evidence [71] and moving towards digital education governance [72]. The knowledge generated here can provide valuable insights for the development of public policies aimed at promoting equity in education. By strategically utilizing this information, it is possible to implement more effective measures and foster a positive transformation in the educational system, benefiting all students and building a fairer and more equitable society. The main idea behind this theory is that by gaining a better understanding of the factors contributing to educational inequality, decision-makers can develop more targeted policies and strategies to promote equity in the educational system. These measures may include allocating additional resources to the most vulnerable schools, implementing specific academic and socio-emotional support programs, and/or adopting pedagogical practices that cater to the individual needs of students.

ACKNOWLEDGMENT

The funders had no role in the study design, data collection and analysis, the decision to publish, or the preparation of the manuscript. Additionally, the funders had no influence on the interpretation of data and the final conclusions drawn.

REFERENCES

[1] M. Ainscow, "Diversity and equity: A global education challenge," *New Zealand J. Educ. Stud.*, vol. 51, no. 2, pp. 143–155, Nov. 2016.

[2] C. Campbell, "Educational equity in Canada: The case of Ontario's strategies and actions to advance excellence and equity for students," *School Leadership Manage.*, vol. 41, nos. 4–5, pp. 409–428, Oct. 2021.

[3] D. D. Ready, V. E. Lee, and K. G. Welner, "Educational equity and school structure: School size, overcrowding, and schools-within-schools," *Teachers College Rec., Voice Scholarship Educ.*, vol. 106, no. 10, pp. 1989–2014, Oct. 2004.

[4] *Monitoring Educational Equity*, National Academies of Sciences, Engineering, and Medicine and Others, National Academies Press, Washington, DC, USA, 2019.

[5] P. J. de los Santos, A.-J. Moreno-Guerrero, J.-A. Marín-Marín, and R. S. Costa, "The term equity in education: A literature review with scientific mapping in web of science," *Int. J. Environ. Res. Public Health*, vol. 17, no. 10, p. 3526, May 2020.

[6] O. Espinoza, "Solving the equity–equality conceptual dilemma: A new model for analysis of the educational process," *Educ. Res.*, vol. 49, no. 4, pp. 343–363, Dec. 2007.

[7] P. E. Poekert, S. Swaffield, E. K. Demir, and S. A. Wright, "Leadership for professional learning towards educational equity: A systematic literature review," *Prof. Develop. Educ.*, vol. 46, no. 4, pp. 541–562, Aug. 2020.

[8] B. Alcott, P. Rose, R. Sabates, and R. Torres, *Handbook on Measuring Equity in Education*, document No. TD/TNC 133.619, UNESCO Institute for Statistics, 2018.

[9] T. N. Hopfenbeck, J. Lenkeit, Y. El Masri, K. Cantrell, J. Ryan, and J.-A. Baird, "Lessons learned from PISA: A systematic review of peer-reviewed articles on the programme for international student assessment," *Scandin. J. Educ. Res.*, vol. 62, no. 3, pp. 333–353, May 2018.

[10] A. L. Sønderlund, E. Hughes, and J. Smith, "The efficacy of learning analytics interventions in higher education: A systematic review," *Brit. J. Educ. Technol.*, vol. 50, no. 5, pp. 2594–2618, Sep. 2019.

[11] P. Perchinunno, M. Bilancia, and D. Vitale, "A statistical analysis of factors affecting higher education dropouts," *Social Indicators Res.*, vol. 156, nos. 2–3, pp. 341–362, Aug. 2021.

[12] C. Herodotou, B. Rienties, B. Verdin, and A. Boroowa, "Predictive learning analytics 'at scale': Guidelines to successful implementation in higher education," *J. Learn. Anal.*, vol. 6, no. 1, pp. 85–95, Apr. 2019.

[13] D. Nusche, "Assessment of learning outcomes in higher education," in *Proc. OECD Educ. Working Papers*, 2008, no. 15, pp. 1–49.

[14] E. Aizikovitsh-Udi and D. Cheng, "Developing critical thinking skills from dispositions to abilities: Mathematics education from early childhood to high school," *Creative Educ.*, vol. 6, no. 4, pp. 455–462, 2015.

[15] M. P. Marchbanks III, J. J. Blake, E. A. Booth, D. Carmichael, A. L. Seibert, and T. Fabelo, "The economic effects of exclusionary discipline on grade retention and high school dropout," in *Closing the School Discipline Gap: Equitable Remedies for Excessive Exclusion*. New York, NY, USA: Teachers College Press, 2015, pp. 59–74.

[16] Y. Lee and J. Choi, "A review of online course dropout research: Implications for practice and future research," *Educ. Technol. Res. Develop.*, vol. 59, no. 5, pp. 593–618, Oct. 2011.

[17] C. Weiland, R. Unterman, A. Shapiro, S. Staszak, S. Rochester, and E. Martin, "The effects of enrolling in oversubscribed prekindergarten programs through third grade," *Child Develop.*, vol. 91, no. 5, pp. 1401–1422, Sep. 2020.

[18] H. Aldowah, H. Al-Samarraie, and W. M. Fauzy, "Educational data mining and learning analytics for 21st century higher education: A review and synthesis," *Telematics Informat.*, vol. 37, pp. 13–49, Apr. 2019.

[19] *Global Citizenship Education: Preparing Learners for the Challenge of the 21st Century*, UNESCO, United Nations Educational, Scientific and Cultural Organization (UNESCO), Paris, France, 2013.

[20] *A Fair Adjustment: Efficiency and Equity of Public Spending in Brazil*, Brazil Public Expenditure Review, World Bank Group, Washington, DC, USA, 2017, vol. 1.

[21] E. M. Queiroga, M. F. B. Machado, V. R. Paragarino, T. T. Primo, and C. Cechinel, "Early prediction of at-risk students in secondary education: A countrywide K-12 learning analytics initiative in Uruguay," *Information*, vol. 13, no. 9, p. 401, Aug. 2022.

[22] A. Suominen and A. Hajikhani, "Research themes in big data analytics for policymaking: Insights from a mixed-methods systematic literature review," *Policy Internet*, vol. 13, pp. 464–484, Jun. 2021.

[23] G. Azzone, "Big data and public policies: Opportunities and challenges," *Statist. Probab. Lett.*, vol. 136, pp. 116–120, May 2018.

[24] S. Capogna, "Sociology between big data and research frontiers, a challenge for educational policies and skills," *Qual. Quantity*, vol. 57, no. 1, pp. 193–212, Feb. 2023.

[25] J. A. Marsh, J. S. McCombs, and F. Martorell, "How instructional coaches support data-driven decision making: Policy implementation and effects in Florida middle schools," *Educ. Policy*, vol. 24, no. 6, pp. 872–907, Nov. 2010.

[26] Y. Wang, "Big opportunities and big concerns of big data in education," *TechTrends*, vol. 60, no. 4, pp. 381–384, Jul. 2016.

[27] A. E. King, F. A. E. McQuarrie, and S. M. Brigham, "Exploring the relationship between student success and participation in extracurricular activities," *SCHOLE, J. Leisure Stud. Recreation Educ.*, vol. 36, nos. 1–2, pp. 42–58, Jul. 2021.

[28] M. Mourshed, M. Krawitz, and E. Dorn, *How to Improve student Educational Outcomes: New Insights From Data Analytics*. New York, NY, USA: McKinsey & Company, Sep. 2017.

[29] A. D. Ewbank, "Student privacy in the age of big data," *Knowl. Quest*, vol. 44, no. 4, pp. 6–7, 2016.

[30] T. Payton and T. Claypoole, *Privacy Age Big Data: Recognizing Threats, Defending Your Rights, Protecting Your Family*. Lanham, MD, USA: Rowman & Littlefield, 2023.

[31] W. Wijayanti, A. N. Hidayanto, N. Wilantika, I. R. Adawati, and S. B. Yudhoatmojo, "Data quality assessment on higher education: A case study of institute of statistics," in *Proc. Int. Seminar Res. Inf. Technol. Intell. Syst. (ISRITI)*, Nov. 2018, pp. 231–236.

[32] S. Gorard, B. H. See, and N. Siddiqui, "What is the evidence on the best way to get evidence into use in education?" *Rev. Educ.*, vol. 8, no. 2, pp. 570–610, Jun. 2020.

[33] E. Zeide, "The structural consequences of big data-driven education," *Big Data*, vol. 5, no. 2, pp. 164–172, Jun. 2017.

[34] C. Baek and T. Doleck, "Educational data mining versus learning analytics: A review of publications from 2015 to 2019," *Interact. Learn. Environments*, vol. 31, no. 6, pp. 3828–3850, Aug. 2023.

[35] I. Hilliger, M. Ortiz-Rojas, P. Pesántez-Cabrera, E. Scheihing, Y.-S. Tsai, P. J. Muñoz-Merino, T. Broos, A. Whitelock-Wainwright, and M. Pérez-Sanagustín, "Identifying needs for learning analytics adoption in Latin American universities: A mixed-methods approach," *Internet Higher Educ.*, vol. 45, Apr. 2020, Art. no. 100726.

[36] A. Bonamino and S. Z. Sousa, "Três gerações de avaliação da educação básica no Brasil: Interfaces com o currículo da/na escola," *Educação e Pesquisa*, vol. 38, no. 2, pp. 373–388, Feb. 2012.

[37] R. Cattell, *The Scientific Use of Factor Analysis in Behavioral and Life Sciences*. Cham, Switzerland: Springer, 2012.

[38] R. L. Gorsuch, *Factor Analysis*. Mahwah, NJ, USA: L. Erlbaum Associatesa, 1983.

[39] J. C. Pinto, H. R. Pinto, and L. Faria, "Autorregulação na aprendizagem: um estudo exploratório no ensino superior portuguçãs," *Psicologia da Educação*, vol. 43, pp. 1–12, Dec. 2016.

[40] L. R. Fabrigar, D. T. Wegener, R. C. MacCallum, and E. J. Strahan, "Evaluating the use of exploratory factor analysis in psychological research," *Psychol. Methods*, vol. 4, no. 3, pp. 272–299, 1999.

[41] E. Bertoni, G. Elacqua, L. Marotta, M. Martínez, H. Santos, and S. Soares, "Is school funding unequal in Latin america? A cross-country analysis of interregional disparities in public spending," *Comparative Educ. Rev.*, vol. 67, no. 1, pp. 100–122, Feb. 2023.

[42] F. J. Murillo and M. Román, "School infrastructure and resources do matter: Analysis of the incidence of school resources on the performance of Latin American students," *School Effectiveness School Improvement*, vol. 22, no. 1, pp. 29–50, Mar. 2011.

[43] N. Gómez-Fernández and M. Mediavilla, "Exploring the relationship between information and communication technologies (ICT) and academic performance: A multilevel analysis for Spain," *Socio-Econ. Planning Sci.*, vol. 77, Oct. 2021, Art. no. 101009.

[44] M. S. Bellibas, "Who are the most disadvantaged? Factors associated with the achievement of students with low socio-economic backgrounds," *Educ. Sci., Theory Pract.*, vol. 16, no. 2, pp. 691–710, 2016.

[45] L. O. Picus, S. F. Marion, N. Calvo, and W. J. Glenn, "Understanding the relationship between student achievement and the quality of educational facilities: Evidence from Wyoming," *Peabody J. Educ.*, vol. 80, no. 3, pp. 71–95, Jul. 2005.

[46] D. S. Edwards, "Another one rides the bus: The impact of school transportation on student outcomes in Michigan," *Educ. Finance Policy*, vol. 19, no. 1, pp. 1–31, 2024, doi: 10.1162/edfp_a_00382.

[47] L. M. Hopson, A. D. Lidbe, M. S. Jackson, E. Adanu, X. Li, P. Penmetsa, H. Y. Lee, A. Anderson, C. Obuya, and G. Abura-Meerdink, "Transportation to school and academic outcomes: A systematic review," *Educ. Rev.*, vol. 76, no. 3, pp. 648–668, Apr. 2024.

[48] E. D. Peteros, S. C. Ypil, J. V. de Vera, G. A. Alcantara, M. D. Fulgencio, D. B. Plando, and L. B. Peconcillo Jr., "Effects of school proximity on students' performance in mathematics," *Open J. Social Sci.*, vol. 10, no. 1, pp. 365–376, 2022.

[49] A. H. Bierbaum, A. Karner, and J. M. Barajas, "Toward mobility justice," *J. Amer. Planning Assoc.*, vol. 87, no. 2, pp. 197–210, 2021.

[50] M. Z. M. Tajuddin and I. Rohman, "The impact of broadband penetration on the student performance in primary and secondary school in Malaysia," in *Proc. 11th Int. Conf. Theory Pract. Electron. Governance*, New York, NY, USA, Apr. 2018, pp. 111–117.

[51] J. Tęcza, S. Wallsten, and Y. Lee, "Do broadband subsidies for schools improve students' performance? Evidence from Florida," May 2023. [Online]. Available: https://ssrn.com/abstract=4478782
[52] A. Grimes and W. Townsend, "Effects of (ultra-fast) fibre broadband on student achievement," *Inf. Econ. Policy*, vol. 44, pp. 8–15, Sep. 2018.
[53] C. Cambini, L. Sabatino, and S. Zaccagni, "The faster the better? The effect of ultra-fast broadband on students' performance," Dept. Econ. Center Econ. Behav. Inequality (CEBI), Univ. Copenhagen, Copenhagen, Denmark, CEBI Working Paper Series 21–14, Sep. 2021.
[54] J. Wainer, T. Dwyer, R. S. Dutra, A. Covic, V. B. Magalhães, L. R. R. Ferreira, V. A. Pimenta, and K. Claudio, "Too much computer and Internet use is bad for your grades, especially if you are young and poor: Results from the 2001 Brazilian SAEB," *Comput. Educ.*, vol. 51, no. 4, pp. 1417–1429, Dec. 2008.
[55] N. Badasyan and S. Silva, "The impact of Internet access at home and/or school on students' academic performance in urban areas in Brazil," *Int. J. Educ. Econ. Develop.*, vol. 9, no. 2, p. 149, 2018.
[56] M. Almasi, H. Machumu, and C. Zhu, "Internet use among secondary schools students and its effects on their learning," in *Proc. INTED 11th Int. Technol., Educ. Develop. Conf.*, Mar. 2017, pp. 2379–2390.
[57] S. Maffei, F. Leoni, and B. Villari, "Data-driven anticipatory governance. Emerging scenarios indata for policypractices," *Policy Design Pract.*, vol. 3, no. 2, pp. 123–134, Apr. 2020.
[58] N. Selwyn, "Data points: Exploring data-driven reforms of education," *Brit. J. Sociology Educ.*, vol. 39, no. 5, pp. 733–741, Jul. 2018.
[59] I. Martinez Lunde, "Learning analytics as modes of anticipation: Enacting time in actor-networks," *Scandin. J. Educ. Res.*, vol. 68, no. 2, pp. 218–232, Feb. 2024.
[60] K. Hasa, "Examining the OECD's perspective on AI in education policy: A critical analysis of language and structure in the 'AI and the future of skills' (AIFS) document and its implications for the higher education," Ph.D. thesis, Dept. Educ. Stud., Univ. British Columbia, Vancouver, BC, Canada, 2023.
[61] S. Sellar, "Data infrastructure: A review of expanding accountability systems and large-scale assessments in education," *Discourse, Stud. Cultural Politics Educ.*, vol. 36, no. 5, pp. 765–777, Oct. 2015.
[62] F. Filgueiras and L. Lui, "Designing data governance in brazil: An institutional analysis," *Policy Des. Pract.*, vol. 6, no. 1, pp. 41–56, Jan. 2023.
[63] S. Washington, "An infrastructure for building policy capability—Lessons from practice," *Policy Des. Pract.*, vol. 6, no. 3, pp. 283–298, Jul. 2023.
[64] F. Filgueiras, "Artificial intelligence and education governance," *Educ., Citizenship Social Justice*, Apr. 2023, doi: 10.1177/174619792311606.
[65] B. Williamson, "Policy networks, performance metrics and platform markets: Charting the expanding data infrastructure of higher education," *Brit. J. Educ. Technol.*, vol. 50, no. 6, pp. 2794–2809, Nov. 2019.
[66] K. N. Gulson and S. Sellar, "Emerging data infrastructures and the new topologies of education policy," *Environ. Planning D, Soc. Space*, vol. 37, no. 2, pp. 350–366, Apr. 2019.
[67] B. Lingard, *The Global Education Industry, Data Infrastructures, and the Restructuring of Government School Systems*. Cham, Switzerland: Springer, 2019, pp. 135–155.
[68] A. N. de Vasconcelos, L. A. Freires, G. D. L. Loureto, G. Fortes, J. C. A. da Costa, L. F. F. Torres, I. I. Bittencourt, T. D. Cordeiro, and S. Isotani, "Advancing school dropout early warning systems: The IAFREE relational model for identifying at-risk students," *Frontiers Psychol.*, vol. 14, pp. 1–17, Jul. 2023.
[69] L. Kyriakides, E. Charalambous, H. P. M. Creemers, and A. Dimosthenous, "Improving quality and equity in schools in socially disadvantaged areas," *Educ. Res.*, vol. 61, no. 3, pp. 274–301, Jul. 2019.
[70] C. Chapman and M. Ainscow, "Using research to promote equity within education systems: Possibilities and barriers," *Brit. Educ. Res. J.*, vol. 45, no. 5, pp. 899–917, Oct. 2019.
[71] R. Gorur and B. Arnold, "Governing by dashboard: Reconfiguring education governance in the global south," in *Digital Disruption in Teaching and Testing*. Evanston, IL, USA: Routledge, 2021, pp. 166–181.
[72] B. Lingard, C. Wyatt-Smith, and E. Heck, "Transforming schooling through digital disruption: Big data, policy, teaching, and assessment 1," in *Digital Disruption in Teaching and Testing*. Evanston, IL, USA: Routledge, 2021, pp. 1–44.

EMANUEL MARQUES QUEIROGA received the degree in analysis and development systems from the Catholic University of Pelotas (UCPel), and the master's and Ph.D. degrees in computer science from the Federal University of Pelotas (UFPel). He is currently engaged in research projects aimed at developing and implementing machine learning and learning analytics systems to enhance academic pathways and ensure academic success within the K-12 framework. His research interests and expertise encompass data science, data mining, big data, machine learning, and learning analytics, with a primary focus on the design of methodologies and practical application within the K-12 education systems.

ELTON SARMANHO SIQUEIRA received the bachelor's degree in physics from the Federal Institute of Pará, in 2008, the bachelor's degree in information systems and the master's degree in computer science from the Federal University of Pará, in 2010 and 2012, respectively, and the Ph.D. degree in computer science from the University of Brasilia, in 2022. He is currently an Adjunct Professor with the Federal University of Pará. He has experience in developing educational games and in developing tools that focus on motivation and the learning process. Additionally, he involved on data science, specializing in predictive modeling.

CARLOS DOS SANTOS PORTELA received the bachelor's degree in information systems from the Federal University of Pará (UFPA), in 2009, the master's degree in computer science from the Federal University of Pernambuco (UFPE), in 2012, and the Ph.D. degree in computer science from the Center for Informatics (CIn), UFPE, in 2017. He has academic experience in the area of software engineering, involved on the following research topics: project management (SCRUM), process modeling (SPEM and BPMN), software process quality models (MPS.BR and CMMI-DEV), software process execution, and teaching-learning of software engineering.

THIAGO DAMASCENO CORDEIRO received the bachelor's degree in computer engineering and the master's degree in electrical engineering from the Federal University of Para (UFPA), in 2006 and 2010, respectively, and the Ph.D. degree in electrical engineering from the Federal University of Campina Grande (UFCG), in 2018. He is currently an Associate Professor with the Institute of Computing (IC), Federal University of Alagoas (Ufal), and a Permanent Member with the Graduate Program in Informatics (PPgI), Ufal. He has been involved in the areas of digital signal processing, with an emphasis on electrocardiogram signals, and the application of artificial intelligence techniques for the classification of heart diseases.

IG IBERT BITTENCOURT is currently a Psychologist, a Computer Scientist, and an Associate Professor with the Federal University of Alagoas (UFAL). He is a CNPq DT-1C Scholarship holder. He is a Visiting Researcher with the Harvard Graduate School of Education. His career is dedicated to United Nations Sustainable Development Goals (SDGs), especially SDG4 (goals 4.1, 4.5, 4.6, and 4.7) and SDG5 (goal 5.1). He is devoted to education, involved on the design, development, and experimentation of intelligent educational technologies to promote equitable learning and well-being. He was a recipient of the National Order of Educational Merit Medal.

SEIJI ISOTANI (Senior Member, IEEE) received the Ph.D. degree from Osaka University, Japan. He was a Postdoctoral Researcher with Carnegie Mellon University. He is currently a Visiting Professor and a Scholar with the Harvard Graduate School of Education and a Professor of computer science and learning technology with the University of São Paulo, Brazil. For more than 15 years, he has dedicated his research career to advancing the science of how people learn with interactive/intelligent educational technologies and designing and implementing public policies to ensure that every student receives the personalized support needed to engage in fulfilling and meaningful educational experiences. He is renowned for his work in the fields of gamification in education, intelligent tutoring systems (ITS), and artificial intelligence in education (AIED). Since 2017, he has been a Technical/Scientific Advisor of Brazilian Ministry of Education, designing and implementing public policies related to educational technologies that have significantly influenced over 50 million students nationwide.

RAFAEL FERREIRA MELLO received the Ph.D. degree in computer science. He is currently a Professor with UFRPE and CESAR. Over the past years, he has supported the adoption of learning analytics techniques and tools in various institutions in Brazil. He leads two nationally scoped projects, funded by Brazilian Ministry of Education, aimed at developing 1) natural language processing techniques and learning analysis to enhance student writing productions and 2) learning analytics dashboards to support decision-making regarding the national education plan. Additionally, he has involved on several multinational research projects involving institutional and organizational partners in Europe, Australia, and Latin America. His research interests include learning analytics, generative artificial intelligence, and natural language processing. He has published in leading international journals and conference proceedings in the areas of his research. He has also served as an Assistant Editor and a Reviewer for various journals and conferences, including the International Learning Analytics Knowledge Conference, *Journal of Learning Analytics*, *Computers and Education*, *British Journal of Educational Technology*, and *The Internet and Higher Education*.

ROBERTO MUÑOZ (Member, IEEE) received the master's degree in computer engineering, engineering science, and higher education, and the Ph.D. degree in computer engineering. He is currently a Full Professor and also the Head of the School of Informatics Engineering, Universidad de Valparaíso, Chile. His research interests include computers and education, multimodal learning analytics, human–computer interaction, and health informatics.

CRISTIAN CECHINEL received the Ph.D. degree in information and knowledge engineering from the University of Alcalá, Spain. He is currently an Associate Professor with the Sciences, Technologies and Health Center, Federal University of Santa Catarina (Brazil). His research interests include the development and analysis of digital learning technologies, learning analytics, and distance learning.