SATPE: Analysis of the Predictive System for the Prevention of School Dropout in the State of Guanajuato, Mexico
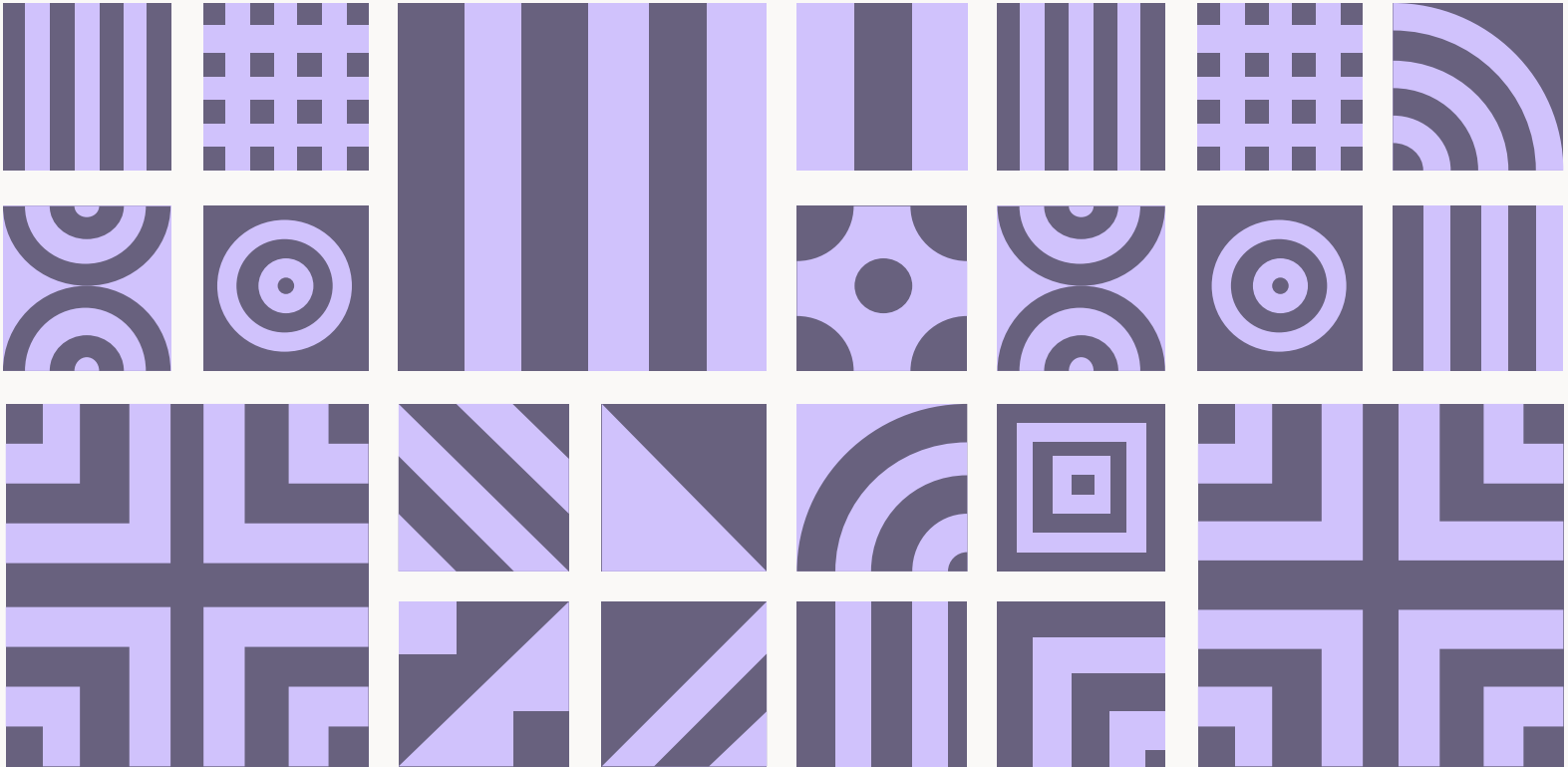
DERECHOS DIGITALES América Latina

SATPE: ANALYSIS OF THE PREDICTIVE SYSTEM FOR THE PREVENTION OF SCHOOL DROPOUT IN THE STATE OF GUANAJUATO, MEXICO

ANALYSIS OF THE PREDICTIVE SYSTEM FOR THE PREVENTION OF SCHOOL DROPOUT IN THE STATE OF GUANAJUATO, MEXICO

The Mexican education system faces steep challenges. On the one hand, despite having promoted numerous initiatives and substantial investments to close gaps in education — which widened during the pandemic — the Mexican government's ability to offer high-quality services is limited in terms of infrastructure, human resources and teaching models, making it difficult to meet the population's needs. On the other hand, the systemic problems afflicting the country — deep inequality and a context of violence — seriously curb access to minimum conditions for students to attend school and complete their education. Finally, given these adverse conditions and taking into account the many variables that come into play at individual, family, academic and community levels, the results of the educational process are unsatisfactory, and academic performance is low (Martinez et al., 2022). In short, given the gravity of the situation, education requires priority attention from society as a whole.

The National Education System in Mexico comprises Basic, Upper Intermediate and Higher types, in school, non-school and mixed formats. Basic education is composed of three levels: Preschool, Primary and Junior High School.[1] Of the myriad challenges facing the system, school dropout is one of most serious socio-educational problems, with the greatest social and economic impact (Vizcaíno et al., 2020). Nationally, the highest incidence of school abandonment is at the upper intermediate level (INEGI, 2023). Starting in 2012 compulsory upper intermediate education was decreed in Mexico to expand the coverage of education services (Arellano & Ortiz, 2022). For the 2022–2023 school year, the dropout rate at the upper intermediate level was 8.7% (INEGI, 2023). Thus, education policies at federal and local levels have been oriented to preventing dropout in junior high and high school.

School dropout, defined as the failure to complete the compulsory school program, is a complex phenomenon encompassing several factors. The National Institute for Statistics and Geography (INEGI) tracks the causes leading young people aged 15–24 to abandon

---

1 Secretaría de Educación Pública. (2015) "Conoce el Sistema Educativo Nacional." Secretaría de Educación Pública, Gobierno de México https://www.gob.mx/sep/articulos/conoce-el-sistema-educativo-nacional (reviewed July 2024).

or drop out of the National Education System based on the following classification: 1) I completed my studies,[2] 2) I no longer wanted to study, 3) I had to work, 4) the school is too far from my home, 5) I had to help with housework, 6) I got pregnant, 7) I got sick, 8) I never went to school. While it is true that this classification includes some of the secondary causes due to which young people leave school, it excludes many other variables involved in the decision to drop out. These may be exogenous or endogenous, such as precarious situations, lack of family or school conditions, and violence in the area, family or school.

Faced with this scenario, all levels of government (federal, state and municipal) have advanced diverse policies and strategies to reduce the percentage of school dropout at the different levels of education. Public policy interventions, mainly promoted by the federal government and local administrations, include a range of actions targeting both schools and individuals. Noteworthy initiatives in schools include fostering the independence of school management, implementing full-time schools, interventions geared toward the development of socio-emotional skills, early warning systems and tutoring programs. For individuals, actions such as funds transfers, differentiated scholarships and teen pregnancy prevention programs are being implemented (Arellano & Ortiz, 2022).

In recent years, prior to the pandemic, these diverse initiatives led to a steady decrease. As of the pandemic, however, the results for reducing dropout urgently need to be addressed to recover the downward trend.

GUANAJUATO STATE

Article 4 of the Guanajuato State Education Act[3] mentions that in the state, "Every person has the right to receive high-quality education in equal conditions" and that "The state and municipalities must provide education services that guarantee learners' highest academic achievement, so that all may pursue preschool, primary, junior high school and upper intermediate education."[4] To achieve this

---

2 It is noteworthy that the classification includes completing school.

3 Congreso del Estado de Guanajuato. Ley de Educación para el Estado de Guanajuato. Last reform: May 29, 2018. https://sep.gob.mx/work/models/sep1/Resource/9db15657-4ea9-47fe-9fe6-4a6181040a2c/guanajuato.pdf (reviewed July 2024)

4 Translator's note: *Todo individuo tiene derecho a recibir educación de calidad en condiciones de equidad* and *El Estado y los municipios están obligados a prestar servicios educativos que garanticen el máximo logro de aprendizaje de los educandos, para que todos puedan cursar la educación preescolar, primaria, secundaria y media superior.*

mandate, the Government of Guanajuato has rolled out a series of education policies and interventions.

The diverse strategies for addressing school dropout and facilitating passage through the education system include projects like "I will stay [*Sí me quedo*]" in high school, "Stay in your high school [*Permanece en tu Prepa*]," and the State Scholarship System. Further, Guanajuato has a Single High School Entrance Exam, combined with an enrollment system for applicants starting in third year of junior high school, whose goal is for all students to be able to enter any of the public upper intermediate educational institutions, even if it is not their first choice. This approach is part of the strategies designed to foster educational continuity (Barquera, 2023, p. 54).

Currently, the Social Contract for Education is the strategy that encompasses these interventions, including *Educational Trajectories*, which involves strengthening mechanisms to foster retention and graduation, using an Early Warning System, and another mechanism related to easing standards for student enrollment and graduation (Barquera, 2023, p. 54). The Trajectories system has been gathering data since 2008 using the Academic Monitoring System in Basic Education. This system makes it possible to record data starting when students enter preschool through their arrival at university. Using this database, a system is developed that includes products such as consulting grades and attendance to notify families, an indicator dashboard for schools, digital records and automatic certification, among others (Barquera, 2023, p. 54).

One of the educational interventions under the Educational Trajectories program is the **Early Action System for School Permanence (SATPE)**, a predictive system used in public schools in the state to identify the risk of dropout among junior high school students. The SATPE program was initially deployed in two stages: a prototype stage during the 2020–2021 school year for a sample of public junior high schools in the state and, later, a pilot phase from 2022 to 2023 in which the system was refined and new databases were incorporated. Currently, the program is operational for all public junior high schools.

THIS PROJECT

To explore the diverse processes, actors, contexts and technologies that come into play in the deployment of these systems, based on analysis of documentary sources and in-depth interviews with various parties involved in the process, the goal of this study is to analyze the implementation of the SATPE program in Guanajuato. To do this, we systematized its implementation considering the following analytical dimensions: the

context for implementation, the legal and institutional framework, the technological design, and the data and decision-making infrastructure. We sought to answer the question of how the SATPE contributed to achieving the goal for which it was deployed, what lessons can be learned from this kind of initiative, and the obstacles governments face in implementing public policies based on data collection and analysis.

PREDICTIVE SYSTEMS IN SCHOOLS: LITERATURE REVIEW

In the academic literature, the phenomenon of school dropout is approached as a long-term academic disconnection, influenced by three levels of factors that are both within school and extracurricular (Freeman & Simonsen, 2015) and which must be considered holistically to address the problem. These levels are *the individual and their immediate surroundings,* including biological aspects, cognitive and psychological skills, motivation and subjective disposition; *the immediate school and social environment,* which includes the quality of the school and services linked to individual wellbeing; and finally, the macrosocial level, comprising factors that influence the decision to continue or abandon studies, but which escape the possibility of modification (Arellano & Ortiz, 2022, p. 35).

In recent years, educational data processing, often linked to the field of learning analytics, has emerged as a growing trend (Urbina-Najera et al., 2020) adopted by governments to support public policy tools. This data-based approach — occasionally incorporating automated decision-making systems — seeks to provide a framework for problem analysis (Arellano & Garcia, 2022) and to thus facilitate generating evidence of the relevance of interventions geared to mitigating school dropout. Some of these interventions are based on educational data analysis to predict academic performance, the creation of predictive retention models, analysis of behavior profiles and school failure, among others (Urbina-Najera et al., 2020). This trend has also been driven and supported by international organizations like UNESCO and the World Bank (2019).[5]

Early warning systems for predicting school dropout have been incorporated in diverse geographic contexts and educational levels. A study conducted in Spain in higher education applied an algorithm for selecting attributes to identify the most important factors influencing the decision to drop out of school. Decision trees were used to

---

5   Melissa Adelman, Francisco Haimovich Paz, & Enrique Alasino. (2019). "¿Quiénes abandonarán la escuela? Aprovechar los sistemas de datos administrativos para predecir la deserción escolar en Guatemala y Honduras." *Education for Global Development*, Banco Mundial. https://blogs.worldbank.org/es/education/quienes-abandonaran-la-escuela-aprovechar-los-sistemas-de-datos-administrativos-para (reviewed July 2024)

define patterns that could warn of imminent school desertion. Through the attribute selection algorithm, 26 relevant factors were found. Three main factors recognized were the lack of guidance, an appropriate student environment and academic monitoring, while the decision tree yielded seven patterns. These included factors such as student environment, inadequate economic support, experiencing an uncomfortable situation, and career choice placement, among others. The study argues that the application of these algorithms must be adapted to the characteristics of the student population in each region (Urbina-Najera et al., 2020). This result is particularly relevant when cultural, social, educational and school contexts are heterogeneous, as is the case in Latin America and Mexico specifically.

In the Mexican context, Vizcaino et al. (2020) propose an automated system for monitoring and prevention of school dropout. The technological solution to the problem meets the need to decrease dropout rates in the technological high school system, which corresponds to the upper intermediate education level. The automated system monitors three variables in the context of school dropout—school absenteeism, academic performance, and behavior—and also generates communication with parents regarding the factors considered. The system has three modules: a website for institutional student monitoring; a real-time embedded software program that makes it possible to track student presence/absence using facial recognition and Raspberry Pi; and a mobile app to send information to parents. In this case, the starting premise for the system design is based on the fact that ongoing follow-up and monitoring of the factors under consideration and communication with parents serve as dropout mitigation strategies.

These cases show that early warning systems are varied and that their efficacy depends not only on the observed indicators, but also on a broader strategy that takes context into account. Likewise, the technological solution may vary and could include different criteria for evaluating dropout risk, as well as multiple technologies, some more intrusive (such as facial recognition) and others less so, such as predictive infrastructure, models and algorithms that are trained based on different data sources.

METHODOLOGY

To analyze implementation of the Early Action System for School Permanence in the state of Guanajuato, Mexico, this study proposes a methodological design based on a documentary, qualitative and transversal analysis. The proposed methodology for this study connects three primary and secondary sources of information: a) document

review of public reports and press statements on the program; b) transparency requests submitted via Mexico's National Transparency Platform;[6] and c) in-depth interviews with stakeholders involved in the development and implementation of the SATPE.

DOCUMENTARY ANALYSIS

a) Public information
The documentary analysis included publicly available information on the program, national and international press releases, publications on official government websites, technical reports and documentary records from other sources in which the program's nature and stakeholders were detailed.

b) Transparency requests
We submitted transparency requests in the categories of the legal framework within which the project operates, obligations and responsibilities, as well as on its design, architecture and implementation. We understand the filing of requests as an exercise to probe the extent of the government's transparency in a project of this nature, involving the participation of public, private and international cooperation initiatives.

c) In-depth interviews
Three in-depth interviews were conducted with public-and private-sector stakeholders who directly participated in the program's development and implementation. To develop the data collection instrument, the interview guide, we started with the analytical dimensions and operationalization of variables proposed by Derechos Digitales. Text analysis of the interviews was conducted using NVivo software.

DIMENSIONS OF ANALYSIS

The operative-analytical framework proposed by Derechos Digitales comprises five analytical dimensions related to the implementation of Artificial Intelligence (AI) systems: national implementation context, regulatory and institutional context, data infrastructure, decision-making process, and technological design. To structure the analysis, we will address the regulatory and institutional context in different sections and the technological design and data infrastructure in the same section. Below, we present a brief review of the dimensions.

---

6 https://www.plataformadetransparencia.org.mx/

a) National implementation context

The national implementation context contemplates the effectiveness of sociodemographic and technological policies at the country level, analyzing dimensions such as age, sex, population distribution and internet access. Indicators such as age brackets, distribution by sex, migrant population and internet access were assessed, with minimum attributes such as age distribution, percentage of internet access and devices used. The questions focused on understanding the demographic distribution, internet access and connecting devices in the country.

b) Institutional dimension

This dimension examines the institutional infrastructure and international cooperation and agreements in the context of the development and application of AI programs. The presence of government institutions for technological development is investigated, along with their scope in areas such as science and telecommunications. Likewise, potential links to international financial and technical cooperation for the implementation of AI programs are explored, as well as the existence of international commitments in the area of human rights protection that could influence the context of the AI investigation.

c) Infrastructure dimension

This dimension focuses on the data infrastructure needed for implementing the AI. The origin of the data used, their access and confidentiality, as well as consent for their use are investigated. Analysis is conducted into whether variables such as gender, age, ethnicity, disability, income and immigration are taken into consideration in the data used. Active and passive transparency mechanisms are examined, as well as accountability in data processing.

d) Technological design dimension

This dimension focuses on the technological design of AI systems. The implemented solution was examined to see if it is original or an adaptation, who developed the algorithm and if it is open source or proprietary. The conditions for procuring the system, the obligations of the parties contracted, evaluation of human rights impact, records of abuse or restrictions, and the technical and impact evaluation of the AI implementation were all investigated.

e) Regulatory dimension

This dimension analyzes the legal and regulatory framework within which the AI is implemented. Issues such as the incorporation of AI in local laws, the existence of specific laws on personal data protection and access to public information, as well as the presence of a national AI strategy were examined. The existence of an agency or

institution charged with overseeing AI initiatives, legislative reforms undertaken for its implementation and the existence of codes of ethical conduct were also considered. In addition, the complaint mechanisms available to those likely to be affected and the accountability associated with the AI implementation were evaluated.

f) Decision-making dimension

This topic addresses the decision-making process for implementation of the AI. The problems AI attempts to solve, alternatives considered, stated objectives and timeframes for execution were studied. Citizen engagement in the decision-making process is evaluated, along with training of personnel, technical evaluation criteria and financing. The participation of external stakeholders, such as international agencies and private enterprise, is also analyzed.

ANALYSIS

Early Action System for School Permanence (SATPE)

On May 4, 2023, the Guanajuato Secretariat of Education held a press conference to present progress in implementation of the SATPE. This project arose from cooperation between the Guanajuato State Government and the World Bank, under the Social Contract for Education and against the backdrop of the Trajectories program.

The SATPE's main objective is to prevent dropout of upper intermediate school students in Guanajuato via the deployment of a prediction, warning and early intervention system. The system is implemented on the premise that it will make it possible to reduce the school dropout rate in the State of Guanajuato. For this program, the Guanajuato government stated that there is no legal instrument signed between the World Bank and the Guanajuato State government, since only technical assistance is involved.

1. GUANAJUATO: NATIONAL IMPLEMENTATION CONTEXT

a) Economic and social context

Guanajuato is a state located in the Central-North region of Mexico. As of 2020, it had a population of 6,166,934 inhabitants (4.9% of the country's total population), with 48.6% men and 51.4% women. The urban population represents 72%, and rural, 28%.[7]

---

[7] Gobierno de México (2024). "Guanajuato. Entidad Federativa." *Data México.* https://www.economia.gob.mx/datamexico/es/profile/geo/guanajuato-gt?redirect=true&schoolAttendance=range3Educ (reviewed July 2024)

In economic terms, US$31.754 million was generated in exports in 2023. For fourth quarter 2023, the Economically Active Population (EAP) was 2,891,928 people. In that same period, the unemployment rate was 3.06% and the informal labor rate was 53%. Economic sectors include agriculture, cattle ranching, mining and manufacturing industries. There is an industrial corridor between the municipalities of Leon, Celaya, Salamanca, Irapuato, Apaseo el Grande, Silao and Villagran. The most noteworthy agricultural and industrial activities in the state are concentrated there.

Violence and homicide rates in Guanajuato have been recognized as the highest in Mexico. In 2019, 44.95 murders per 100,000 inhabitants were recorded.[8] From January to September 2023 there were 2,024 murders.[9] The number of disappeared people is 3,666.[10]

In terms of emigrating population, the National Institute for Statistics and Geography (INEGI) records that from 2015 to 2020, 92,215 people left Guanajuato to settle in another state.[11] Regarding access to technology, 48.2% of households have internet access and 34.7% have a computer.

b) Educational context

Considering all levels of education, in Guanajuato there are a total of 10,744 schools, 82,409 teachers and an enrollment of 1,685,641 students.[12],[13] Guanajuato's illiteracy

---

8   Arturo Ángel (2020). "Con 35 mil 588 asesinatos, 2019 es el año más violento del que se tenga registro." *Animal Político*. https://www.animalpolitico.com/2020/01/homicidios-2019-violencia-asesinatos-record (reviewed July 2024)

9   Redacción El Economista (November 2023). "Guanajuato es la entidad con más homicidios dolosos del país." *El Economista*. https://www.eleconomista.com.mx/politica/Guanajuato-es-la-entidad-con-mas-homicidios-dolosos-del-pais-20231104-0008.html (reviewed July 2024)

10 Marcela Nochebuena (May 2023). "Con seis asesinatos desde 2020, Guanajuato es la entidad más violenta para personas buscadoras; grupos de madres exigen protección." *Animal Político*. https://www.animalpolitico.com/sociedad/guanajuato-asesinatos-entidad-violenta-personas-buscadoras (reviewed July 2024)

11 INEGI (n.d.). "Movimientos Migratorios." *Cuentame INEGI*. https://cuentame.inegi.org.mx/monografias/informacion/gto/poblacion/m_migratorios.aspx?tema=me&e=11 (reviewed July 2024).

12 Secretaría de Educación del Estado de Guanajuato. (n.d.). "Indicadores Educativos." (available July 2024) http://seg-qlik02.seg.guanajuato.gob.mx/QvAJAXZfc/opendoc.htm?document=prueba%5cindicadores%20educativos.qvw&lang=en-US&host=QVS%40qlikview02&anonymous=true (reviewed July 2024)

13 In terms of junior high school, there are 1,628 schools and 14,017 teachers serving 309,998 students, of whom 154,677 are women and 155,321 are men.

rate in 2020 was 5.27%. Of the total illiterate population, 40.9% corresponds to men and 59.1% to women.[14] The Institute for Adult Literacy and Basic Education (INAEBA) documents that as of 2021, there were around 192,000 people over the age of 15 in Guanajuato who are illiterate.[15]

Although rates of school dropout vary by educational level and state, Guanajuato is above the national average and is positioned among the three states with the highest percentage of school dropout in junior high school (3.1%) and upper intermediate education (10.9) for the 2022–2023 period.

c) Public policy context: the Social Contract for Education

The pandemic contributed to worsening the school dropout rate. In 2019, there were an estimated 65,000 students outside the school system.[16] During the pandemic, although official figures recognize that from 2020 to 2022 around 75,000 students stopped studying, other figures indicate that school dropout could have risen to as many as 100,000 students.[17]

In this context, as an initiative to address the pandemic's effects on education,[18] the Guanajuato government launched the *Social Contract for Education* with support from the World Bank and the UNESCO Mexico Education program and an investment of 1 billion MXN. Of this, 400 million MXN[19] were allocated to recovery of learning (a delay of 1.5 years was estimated due to the pandemic); recovery of school attendance; recognition of the role and work of teachers; and preparing 21st century parents

---

14 Gobierno de México. (2024). "Empleo y Educación." In *Guanajuato. Entidad Federativa*. DataMéxico. https://www.economia.gob.mx/datamexico/es/profile/geo/guanajuato-gt?redirect=true#education-and-employment. (reviewed July 2024).

15 Serna, Patricio. (October 12, 2022). "Tiene Guanajuato más de 190 mil personas analfabetas." *Zona Franca*. https://zonafranca.mx/politica-sociedad/educacion/tiene-guanajuato-mas-de-190-mil-personas (reviewed July 2024).

16 Gasca Ramírez, Yajaira. (2022). "Educación se Volvió un Lujo: 100 Mil Alumnos Fuera de las Aulas en Guanajuato." *PopLab*. https://poplab.mx/v2/story/Educacion-se-volvio-un-lujo:-100-mil-alumnos-fuera-de-las-aulas-en-Guanajuato (reviewed July 2024).

17 Gasca Ramírez, "Educación se Volvió un Lujo."

18 Secretaría de Educación del Estado de Guanajuato. Pacto Social por la Educación. https://www.seg.guanajuato.gob.mx/PactoRegresa/SitePages/Inicio.aspx

19 Unless otherwise indicated, the amounts are expressed in Mexican pesos.

through greater family participation in education, as well as promoting school harmony. Another 600 million MXN were allocated to investment in education infrastructure.[20] The strategy proposed collaboration among various stakeholders, including municipal authorities and private enterprise.[21]

Under the *Social Contract for Education*, the Guanajuato Secretariat of Education considered five causes of school dropout: grade repetition, lack of family support in school life, lack of prevention of risky behavior, low interest in school due to belief that it is unprofitable and, in the case of boys, cultural expectations that they must be providers.[22] The strategy incorporates actions for identifying children and young people who have abandoned their studies, so that schools may promote their reintegration and offer follow up.

The Social Contract for Education is structured in four main components. First, ensuring that all students attend school, which means recovering, reintegrating, and re-enrolling students. Second, guaranteeing that no-one is left behind, focusing on recovery of learning. Third, social recognition of the role of teachers, highlighting their importance in the educational process. Fourth, involving families through "Preparing 21st Century Parents." In addition, a crosscutting strategy is included targeting promotion of school harmony (Barquera, 2023, p. 55).

Some critics consider that the initiatives proposed by the *Social Contract for Education* fail to address the root causes behind school dropout. For example, the causes of dropout exclude economic causes, family violence and mental health, which despite existing prior to the pandemic, were exacerbated in that context.[23]

Regarding the economic variable, poverty in Guanajuato grew by at least one percentage point between 2018 and 2020, rising from 41.5% in 2018 to 42.7% in 2020.[24] There are studies showing that, based on the increase in poverty, the rate of child labor also increased, which affects dropout. In addition, other variables such as family

---

20 Gasca Ramírez, "Educación se Volvió un Lujo."

21 Gasca Ramírez, "Educación se Volvió un Lujo."

22 Gasca Ramírez, "Educación se Volvió un Lujo."

23 Gasca Ramírez, "Educación se Volvió un Lujo."

24 Gasca Ramírez, "Educación se Volvió un Lujo."

displacement due to the context of violence in the state must also be considered a relevant factor.[25]

Despite the multiple interventions implemented by public education policy, the pandemic's effects on the percentage of school dropout have not yet been vanquished (Arellano & Ortiz, 2022). The state government asserts that of the students who had dropped out of school between 2020 and 2022, only 13,000 have been re-integrated into the school system.[26]

2. INSTITUTIONAL DIMENSION: ORGANIZATIONAL STRUCTURE AND STAKEHOLDERS

Institutional context, implementation and benchmarks

The SATPE is one of the Educational Trajectories program initiatives coming under the Office for Educational Innovation in the Mentefactura program[27] and special projects directly managed by the Guanajuato government's Secretary of Education. Given the institutional complexity of the Secretariat of Public Education, this institutional structure addresses the need to find flexible mechanisms and organizational charts (Fig. 1) that facilitate incorporating technological innovation.[28] Within the Social Contract for Education, the Early Warning System is part of the re-integration component. This system attempts to ensure that students not only return to school but also stay there.

---

25 Gasca Ramírez, "Educación se Volvió un Lujo."

26 Gasca Ramírez, "Educación se Volvió un Lujo."

27 Mentefactura is a public policy that attempts to position Guanajuato as a national center for innovation and entrepreneurship. https://www.stratego-st.com/mexico-manufacturing/mentefactura/

28 As per the interview with the head of the Office for Educational Innovation.

Secretaría de Educación del Estado de Guanajuato

Secretario de Educación

Mentefactura y proyectos especiales

Dirección de Innovación Educativa

Trayectorias Educativas

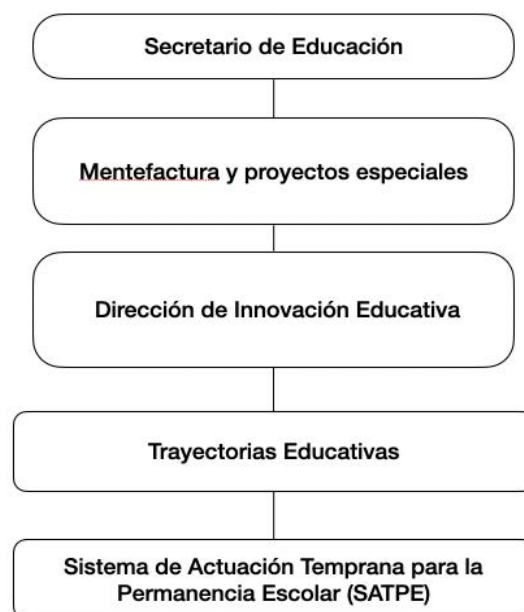Sistema de Actuación Temprana para la Permanencia Escolar (SATPE)

Figure 1. Institutional context for implementing the Early Action System for School Permanence

The SATPE has been implemented for two school years. The first year, 2022–2023, was the prototype stage. During this stage only half (671) of public junior high schools in the state were part of the implementation, with one group of control schools and another of intervention schools. In early 2023 the first list of at-risk students was identified. In the second school year, 2023–2024, all schools were incorporated (1,342), new datasets were integrated, and the model was fine tuned. The program is in a new implementation cycle whose results will be published by the corresponding agencies at the end of the 2023–2024 school year.

The Early Warning System comprises an action framework based on school management as a collective. The objective of the Action Framework for School Permanence (MAPE) is to identify students who are at risk of dropping out of school so that, based on this identification, actions may be initiated by the academic community (Barquera, 2023, p. 56). The action framework makes it possible to define actions that will guide intervention in the school system once results are obtained from the program. Thus, we find that the SATPE system's two components are: 1) the early warning predictive system (algorithm) and 2) the Action Framework for School Permanence (Fig. 2).

Figure 2. SATPE program components

Project milestones include learning analytics, algorithm construction, action framework construction, strategy implementation and evaluation of the strategy (Fig. 3) with the participation of diverse stakeholders at each key moment.

Figure 3. Project milestones

Stakeholders

As a result of our study, we found that 10 organizational entities participated in the SATPE's development and implementation, including public and private agencies, international organizations and government agencies, in addition to the 1,342 public junior high schools in Guanajuato (Fig. 4). The roles of these stakeholders in the project are described below.

- **1,342 secondary schools.** The Guanajuato Secretariat of Education provided for 1,342 schools in the pilot version, including diverse educational models such as

general junior high schools, technical schools and tele-secondary schools that correspond to the entirety of junior high schools in that agency.

- **World Bank.** The project acknowledges the World Bank as the organization responsible for preparing the development and implementation documentation and for providing technical assistance in both the construction and evaluation of the predictive model.

- **Antenalabs.** This is a private information technology services and consulting company. In the project it is recognized as the party responsible for providing consulting through development of the predictive model.

- **PIT Policy Lab.** This is a private organization offering consulting services on public policy linked to digital technologies with a gender-based dimension. In the project it is acknowledged as being responsible for coordinating development of the *Preventing and mitigating gender bias in AI-based early alert systems document.*

- **USAID.** The project recognizes USAID as the agency that provided financial resources for developing the recommendations. Preventing and mitigating gender bias in AI-based early alert systems for education to be developed by Tad, PIT Policy Lab, Women in Digital Transformation and Athena Infonomics.

- **Nutanix.** This is a private company providing cloud-based computing services. They offer infrastructure rental in data centers for applications. In the project it is recognized as the company hired to host the SATPE software applications.

- **Microsoft.** This is a multinational software production and licensing company. In the project it is acknowledged as the company providing the Power Bi service, a platform used for handling project data. The Integrated Educational Information System housing all the state's educational data is hosted on this platform.[29]

- **Gama Sistemas S.A. de CV.** This is the provider responsible for the purchase and sale of the Microsoft Power Bi Software platform user license. The platform is used for processing data related to educational indicators and school monitoring implemented by the SATPE via public tender under contract DGRMSG-119-2023.

---

29 Secretaría de Educación del Estado de Guanajuato. (2024). "SIIE. Sistema Integral de Información Educativa." https://www.seg.guanajuato.gob.mx/SIIE/SitePages/SIIE.aspx (reviewed July 2024)

- **Keymax Comercializadora.** This is the supplier in charge of procuring Nutanix services for the Guanajuato Secretariat of Education via public tender under contract DGRMSG-071-2022.
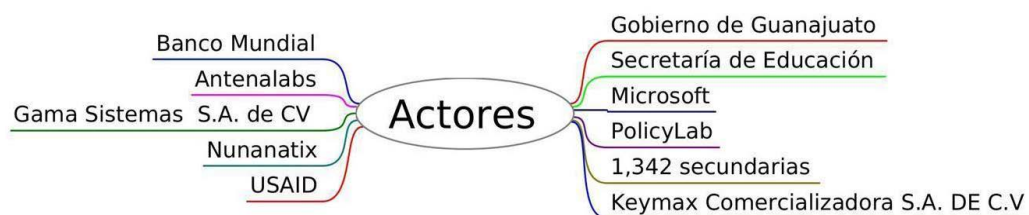

Figure 4. Stakeholders involved in SATPE implementation

3. TECHNOLOGICAL AND DATA INFRASTRUCTURE DIMENSION

Information system architecture

According to the World Bank and the Guanajuato State Secretariat of Public Education, the goal of the "Early Action System for School Permanence" project is to prevent secondary school students in Guanajuato from dropping out of school. For this goal it proposes, in its own words: to identify students at risk of dropping out and take action to avoid their desertion. The system proposes two tools: a **list of students at highest risk** and an **action framework.**

Data sources

This information system uses data collected by the School Monitoring System, the Official State Catalog of Schools (CEO), the Data Collection for Improving Learning Outcomes (RIMA) survey, and information regarding teaching staff. Likewise, data collected by the Guanajuato Secretariat of Public Education prior to the project's development are incorporated (Fig. 6).

Figure 6. Integration of SATPE system databases

Predictive model

The SATPE system makes use of a machine learning model based on historical data mentioned above. For this purpose, the learning model known as XGBoost was chosen.

XGBoost

XGBoost is an open-source machine learning model. It was developed as the result of a research project by Tianqi Chen (https://github.com/tqchen/xgboost) of Carnegie Mellon University. It rose to prominence in data science in 2020. This fame was influenced by its success in the Higgs Boson Challenge organized by Kaggel, now owned by Google. The model is configured as a library in different programming languages such as Python, R, and C/C++, among others (Kunapuli, 2023).

Outcomes

The expected outcome of this learning algorithm is a list of students categorized based on their risk of school dropout. The analysis of 16 variables is conducted using the predictive model offering a percentage of risk translated into a traffic light code (red=highest risk, yellow=some risk, green=no risk).

List of most at-risk students

To build the list of most at-risk students, an information system was developed using an XGBOOST-type supervised machine learning method for classification and regression, as mentioned above (Fig. 7).
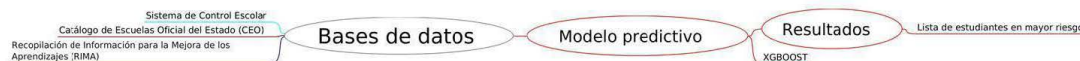
Figure 7. SATPE information architecture

According to information obtained from official sources, the SATPE system is, therefore, a predictive system that makes it possible to generate lists of students at risk of school dropout. The model comprises four stages: data extraction, analysis, predictive model generation and final list generation (Fig. 8).
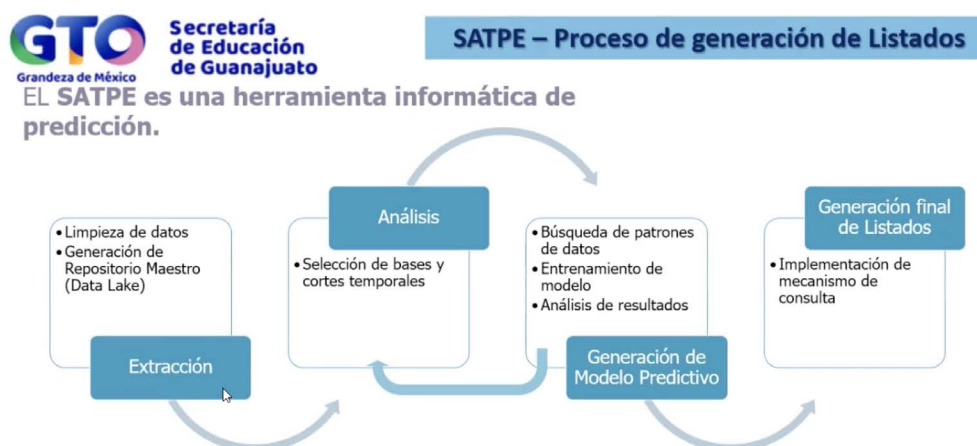

Figure 8. SATPE model stages

System components

According to the requests for information filed, the SATPE is an information system with the following components:

- The computing resources used by the SATPE system belong to the computing service on the Nunatix cloud contracted by the Keymax Comercializadora S.A. de C.V. intermediary, under public tender at a cost of two million eighty-six thousand eight hundred thirty-six (2,086,836) MXN.

- The databases used by the SATPE were built (according to the World Bank).

- The SATPE includes the use of data collected via the Scholastic Records System, the Official State Catalog of Schools (CEO), the Data Collection for Improving Learning Outcomes (RIMA) survey and information on teaching staff.

- Data on educational indicators and scholastic records implemented for the SATPE is processed using Power Bi Software, sold by the Gama Sistemas S.A. de C.V. intermediary under public tender at a cost of twelve million three hundred forty-one thousand four hundred seventy-two (12,341,472) MXN.

- A predictive model developed by Antenalabs for the prototype phase.

- The World Bank conducted the evaluation and systematization of development and implementation.

- One thousand three hundred forty-two (1,342) secondary schools participate in the project.

4. REGULATORY DIMENSION

Legal operating framework and regulatory context

Bills have been filed in Mexico proposing the establishment of a legal framework on the use and development of Artificial Intelligence systems. The most recent is the Ethical Regulation of Artificial Intelligence and Robotics Act[30] submitted in May 2023. This bill proposes the build-up of an agency that could create official standards for Artificial Intelligence systems. It also suggests that this same agency could propose regulations on the topic, in addition to official standards (González, 2023).

Data protection and applicable laws

The SATPE in its prototype stage and its current phase is implemented under the framework of public education institutions. Therefore, in terms of data protection, the Secretariat of Public Education is the agency responsible for collecting personal data, its processing and the protections offered it, as well as for complying with what are known as Access, Rectification, Cancellation and Opposition (ARCO) rights enshrined

---

30 Loyola Vera, Ignacio. (March 2023). "Iniciativa que Expide la Ley para la Regulación Ética de la Inteligencia Artificial para los Estados Unidos Mexicanos, Suscrita por el Diputado Ignacio Loyola Vera y Legisladores Integrantes del Grupo Parlamentario del PAN." Sistema de Información Legislativa de la Secretaría de Gobernación. http://sil.gobernacion.gob.mx/Archivos/Documentos/2023/04/asun_4543395 _20230413_1680209417.pdf.

in the Protection of Personal Data in the Possession of Obliged Subjects Act for Guanajuato State.

The data used by the SATPE correspond to a dataset previously built for the purposes of the Scholastic Records System, the Official State Catalog of Schools, the Data Collection for Improving Learning Outcomes survey and teaching staff. These data were collected using the Simplified Privacy Warning[31] of the Office of School Services through which users (parents and tutors) are informed of the purposes for which information on children and adolescents is collected, obtaining their tacit and — in some cases — express consent, according to statements from the Transparency Unit of the Guanajuato State Executive Branch.

Given that the data were collected under the School Monitoring System, the Official State Catalog of Schools, and the Data Collection for Improving Learning Outcomes survey, we investigated data protection considerations at the time the pilot and current phases of the SATPE were implemented. Our analysis in this regard was to become familiar with the considerations and grounds for authorization of the Secretariat of Public Education in terms of using these data for the SATPE system.

The grounds for authorization we found indicate that it was not considered necessary to prepare a different warning from the one that already existed. This is due to the fact that the current warning covers "the use of data to integrate or modify databases in our electronic systems, operating and statistical effects" and further that "the information is kept in the Scholastic Records System."[32]

Indeed, during our analysis we found that this paragraph is present in the Simplified Privacy Warning of the Office of School Services targeting the collection, processing and querying of the three systems mentioned above.

While authorization regarding data collection and processing is covered and resolved in terms of the justification and rationale for the purposes stated in the privacy warning, ARCO mechanisms remain an afterthought in the specific case of the SATPE. It is not covered specifically, but rather as a product of use generally

---

31 Secretaría de Educación del Estado de Guanajuato. (n.d.) "Aviso de Privacidad." (Available July 2024) https://www.seg.guanajuato.gob.mx/SitePages/Aviso_privacidad.aspx.

32 Translator's note: *"incluye el uso de datos para integrar o modificar las bases de datos de nuestros sistemas electrónicos, efectos operativos y estadísticos"* and *"la información se mantiene en el Sistema de Control Escolar"*.

foreseen as "use of data to integrate or modify databases in our electronic systems, operating and statistical effects."

We can thus suggest that, from a rights perspective, the project proposes that the integrated information systems and data protection are at the intersection of rights with the right to education. The rationale of applicable legal precepts, reasoning and grounds for action of the educational authorities justify its scope in compliance with the legal framework of the Secretariat of Public Education's operating standards.

In addition to the Protection of Personal Data in the Possession of Obliged Subjects Act, the data collection comes under the regulations of Articles 3(6)(A)(III) and 16(2) of the Constitution of the United Mexican States; Article 3 of the Constitution of the State of Guanajuato; the General Education Act; the General Teaching Careers System Act; the State of Guanajuato Education Act; Articles 13(III), 15 and 25 of the Organic Law for the Executive Branch for the State of Guanajuato; Articles 3(I), 13, 16, 20, 22, 34, 36, 37, 39, 40, 42, 96, 97, 98, 99, 100 and 101 of the Protection of Personal Data in the Possession of Obliged Subjects Act for Guanajuato State; the Internal Regulations of the Secretariat of Education; and Articles 125 and 126 of the General Guidelines for Human Resources Administration attached to the Ministries and Public State Administration Entities.

Evaluations, results and reports

Currently there are technical reports documenting the project's development and implementation, including a report by the Government of Guanajuato systematizing the evaluation of implementation of the SATPE and a report by the World Bank, in which the development and its implementation are systematized. A third public policy recommendations document prepared by Tad, PIT Policy Lab, Women in Digital Transformation and Athena Infonomics was developed under the Equitable AI initiative based on an intervention to identify gender biases in the program, using USAID funds.[33]

Based on the interviews conducted with public officials and statements on the program, perception of the system's usefulness is positive in terms of raising school administration awareness on the problem of school dropout. While it is true that dropout is a fact, the idea of having data providing evidence of the problem makes it possible to undertake actions enabling intervention to avoid it. The system has

---

33 Itad, PIT Policy Lab, Women in Digital Transformation, and Athena Infonomics. (2023). "Prevenir y Mitigar el Sesgo de Género en los Sistemas de Alerta Temprana Basados en IA para la Educación." https://www.policylab.tech/_files/ugd/0e03be_781265d6003349a1ad45b6b63079042d.pdf?lang=es.

not contributed to reducing school dropout, and the state, along with the Federal government, continues to promote strategies that facilitate addressing a problem that is both complex and multi-factor. Something that also stands out in the testimonies is the conviction that public policies should be based on data (Fig. 9), and in terms of challenges, there is a noteworthy need to collect high-quality data and for systems that are interoperable so that strategies can be more effective.

Figure 9. Data-based educational policies

CONCLUSIONS

In light of the critical situation facing academic performance, the adverse conditions existing in complex contexts, and the exacerbation stemming from the COVID-19 pandemic, there is a regional trend toward incorporating predictive systems in schools, that are driven both by national and local government and by international organizations. These predictive systems offer a snapshot of dropout risk among students. The deployment of these systems requires complex coordination among stakeholders, processes, laws, data, and technology but also the understanding that the system must be integrated into a broader educational strategy and policy that make it possible to address this multi-cause problem. As analyzed in this case, a broad-based approach must consider action frameworks that go beyond technological solutions or a particular technology. We hope that this analysis offers evidence regarding the real potential for implementing data-based public policies in contexts like that of Mexico or Latin America, and that it may pose questions regarding the role of international and public–private collaborations in these policies.

REFERENCES

Arellano-Esparza, C.A., Ortiz-Espinoza, Á. "Educación media superior en México: abandono escolar y políticas públicas durante la covid-19." Iconos vol. 74 (2022): 10.17141/iconos.74.2022.5292.

Barquera Arteaga, E. I. Seminario: Prevención del Abandono Escolar en la Educación Media Superior y Superior en Oaxaca. Oaxaca: Gobierno del Estado de Oaxaca-Banco Mundial, 2023.

Freeman, J., & Simonsen, B. "Examining the Impact of Policy and Practice Interventions on High School Dropout and School Completion Rates: A Systematic Review of the Literature." *Review of Educational Research* 85, no. 2 (2015): 205-248. https://doi.org/10.3102/0034654314554431.

González, Fernanda. "Presentan Propuesta de Ley para Regular la IA en México." Wired, May 31, 2023. https://es.wired.com/articulos/diputado-presenta-propuesta-de-ley-para-regula-la-ia-en-mexico (consulted July 2024)

Instituto Nacional de Estadística y Geografía (INEGI). "Tasa de abandono escolar por entidad federativa según nivel educativo, ciclos escolares seleccionados de 2000/2001 a 2022/2023." Instituto Nacional de Estadística y Geografía (INEGI). https://www.inegi.org.mx/app/tabulados/interactivos/?pxq=Educacion_Educacion_11_c6aa7c65-4d89-4eaf-972e-431727fc686d (consulted July 2024)

Urbina-Nájera, A. B., Camino-Hampshire, J. C., & Cruz Barbosa, R. "University Dropout: Prevention Patterns through the Application of Educational Data Mining." *RELIEVE -Revista Electrónica de Investigación y Evaluación Educativa.* https://doi.org/10.7203/relieve.26.1.16061.

Vizcaíno, Y., Durán, O., García, R., & González Dueñas, M. "Arquitectura de un Sistema Automatizado para el Monitoreo y Prevención del Abandono Escolar." In *Applications in Software Engineering -Proceedings of the 9th International Conference on Software Process Improvement, CIMPS 2020.* doi: 10.1109/CIMPS52057.2020.9390097