UNDERSTANDING THE NEW LEARNING LANDSCAPE:
Accelerating Learning Analytics and AI in Education

MARCH 2021
Abstract

Hundreds of millions of learners have shifted to digital learning environments during the current pandemic. Education policymakers and parents around the world are asking how to make remote and hybrid learning work better. The need for learning analytics and AI to accelerate the pace of our understanding has never been more urgent. This paper describes the work of an international collaboration of large education systems and Microsoft Education to address this urgent need by developing and applying learning analytics and AI to inform educational strategies in the current context. We are putting to use the latest cloud and AI technologies in responsible ways that are sensitive to data privacy, security, and the special accountabilities of education systems. Finally, through publications like this and an open-source repository of technical assets, we are sharing our insights and developments with education systems around the world. This is the first in a series of communications, which will include papers on remote learning patterns and outcomes in diverse school contexts, as well as documentation on the processes by which education systems can implement AI systems ethically and responsibly using their own system data.
An AI Vignette for Education Systems

Like countless other school systems around the world in 2020, the Department of Education in "CityState" was worried about students continuing to progress in their learning when schools were closed. Would teachers be able to engage those hard-to-reach students when they were not physically in the same location? How would students with special needs receive extra support? What about those families in which parents were not able to help children with their schoolwork, due to a lack of technical or educational skills, or not speaking the language of instruction?

CityState’s first step was to set up Open Education Analytics to investigate how they could improve remote learning to ensure every learner stayed engaged in learning. With Open Education Analytics in place, analysts were able to develop a machine learning model that could predict the elements most closely related to student learning progress while schools were physically closed. In the case of CityState, these elements were students’ overall digital engagement over the past two weeks, the existence of school-based teacher professional development, and students’ use of curriculum-specific education applications. With these insights, system leaders developed and tested programs for schools and students to optimize student engagement in learning. For example, when system dashboards showed that student engagement was low across an entire school, system leaders worked with school leaders to conduct teacher training targeted to the needs of that school’s teachers.

The system was also able to identify individual students with low or decreasing engagement. For those students, CityState designed a structured support system based on students’ daily activities in their digital learning environments. As soon as there were data signals of a student’s disengagement, the system notified a network of people supporting that student: first their teachers, then parents or guardians, then counsellors, and others as needed. At each stage of support, the individuals in that network made decisions about how to act based on their understanding of the student’s context, keeping human decision-making deeply embedded in the support process. All the individuals involved in this process could also provide input back into the system to improve its analytics. Most students and families chose to participate in this structured support system and explicitly agreed to share their digital engagement data with the system. However, some families chose not to share their data, or to control which student data was shared with the learning analytics system.
The need to accelerate learning analytics and AI

This story paints a picture of what is possible with modern learning analytics and artificial intelligence. When educators at a system, school and classroom level share a vision of how data should be used to improve learning, the impact on learning throughout the system can be significant. Learning analytics and AI can serve many purposes that would not be possible without the use of rich and continuous real-time data:

- Alerting teachers, school leaders, parents and social or community support organisations to students that are at risk of various well-being, academic or social issues, so that support may be provided before problems deepen.ii
- Recognising patterns that signal when some populations of students might be underserved by current educational or support provisions.iii
- Providing continuous feedback on student engagement, motivation, and learning progress that informs teachers’ changes in curriculum, pedagogy, or practice.iv
- Transforming assessment from a post-hoc snapshot of student recall at a particular moment in time to a rich collection of measures and learning artifacts that show students’ unique interests, strengths, weaknesses, and progress over time and provide individualised recommendations for ways to improve.v
- Providing parents with real-time understanding of their child’s progress in much more detail than what periodic report cards or infrequent parent-teacher communications can provide, reporting on a variety of outcomes (well-being signals, academic achievement, creativity, collaboration, participation, and motivation).vi
- Providing personal progress data, reminders, and recommendations directly to students, enabling them to be more self-aware, better self-manage and regulate their own learning, monitor their own progress, and seek support when needed.xii

Until recently, the use of education data has been an exercise in examining static data from the past that resided in separate data silos to inform decisions about an uncertain future. Schools and school systems waited months or even years for the results of high-stakes exams that would tell them if their programmes, professional development, or other investments improved learning outcomes.
At the same time, other industries are making use of real-time data to assess progress against goals and strategies continuously. Retail stores using AI can determine within 24 hours the impact on purchasing behaviors in response to moving the shelf location of a product. Education systems should be able to see the impact on student well-being and learning of changes in teaching strategies, tools, and programs just as quickly, so they can become more proactive in diagnosing and acting in ways that truly support every learner.

Our goal is that every student will achieve his or her full potential—this is why we do everything. All students come from different backgrounds and have different skill sets. The education system has to personalize learning and needs knowledge to be able to address this. Finland isn’t a country rich in oil or other natural resources; we have to build knowledge capital amongst our citizens. Our goal is not to lose any student in the education process.

**Pasi Silander**
Head of ICT Development Programs, AI & Data Analytics, City of Helsinki, Education Division

One of the side effects of the global COVID-19 pandemic is that students around the world quickly transitioned to remote learning. This mass shift to Internet-connected learning provided tremendously more learning data on a huge scale. The Microsoft O365 and the Teams digital learning environment were at the epicentre of this shift for millions of students. Microsoft Education and education systems realized the need to accelerate and democratise the use of this data through learning analytics and AI, given the global need to understand how to improve remote learning.

In response, Microsoft Education developed **Project Constellation**, a collaboration with education systems committed to advancing the use of learning analytics and AI. Project Constellation is a starting point for broad capacity-building in learning analytics across the education sector. The components of the collaboration include:

1. Open-source technical resources, including a standard architecture for education analytics and AI in the cloud, modern data security and privacy tools, data visualisation and reporting templates, and repeatable machine learning models. These resources are being published in an open-source community called “Open Education Analytics,” with education systems contributing to that community as they develop data and AI assets that can be re-used by other education systems.
2. Capacity-building to help education systems develop their own expertise in analytics and AI, including a set of online learning paths for data engineers and data scientists to develop the skills needed to implement the data architecture and conduct learning analytics. These learning paths are already available for education systems and their technology solution partners.

3. Applying principles of responsible data and AI to learning analytics to ensure the appropriate and ethical use of data by education systems and educators. Microsoft’s expertise and services in data security, privacy, governance, management, and AI are all incorporated into the technical resources, capacity-building, and AI process guidance.

4. A community of practice among large education systems, collaborating in developing these resources, defining best practices, and sharing insights from learning analytics. In particular, as insights are developed and applied on how to improve remote and hybrid learning, these will be published as quickly as possible.

The education systems currently collaborating in Project Constellation include the City of Helsinki Education Division (Finland), the Fresno Unified School District (California), the Azerbaijan National Ministry of Education, and the Tasmania Department of Education (Australia). Combined, these systems hold data for more than 1.2 million students, and they are beginning to make use of the daily digital activity data from those students to inform decisions about how to improve learning and support every student during remote learning and beyond. This scale of data has never been seen before in learning analytics. Project Constellation could ignite a new era of learning analytics for primary and secondary education systems. While learning analytics has been practiced in higher education at the course and institutional level, this new effort seeks to analyze patterns and build predictive models across large systems (city, province, municipality, national). It will utilize all the different types of data these systems hold, from student information systems, learning management systems, assessments, surveys, and edtech applications.

As each education system develops its own data environment and learning analytics capabilities, they will each be able to provide students, families, teachers, and schools insights into student engagement and progress in the moment, rather than seeing data a week, a month, or a semester later. Educators will have real-time data about student attendance, participation, behaviour, actions, processes, and engagement that could inform their teaching and help them provide more personalized supports and better learning experiences for individual students, continuously. Data on this scale—captured every day from multiple applications using multiple measures across a system and matched with data from other system sources—are too much for humans to manually process and analyse. They require modern data tools and the process of learning analytics and AI.

At this stage in the Project Constellation journey, we can report on two phases of this work: the importance of developing a shared purpose for learning analytics and how we are implementing AI processes in ways that directly apply ethical and responsible principles.
Data-driven educational change: learning analytics and AI with a purpose

Canadian education leader Michael Fullan speaks about educational change being driven by a moral purpose. Moral purpose includes the beliefs, culture, values, ideas, practices, and behaviour of humans that drives decisions and action. As Fullan says:

*Moral purpose is about both ends and means. In education, an important end is to make a difference in the lives of students. But the means of getting to that end are also crucial.*

Deciding to develop and employ large-scale learning analytics requires policy-makers, school leaders, teachers, parents, students, and the community to subscribe to a shared purpose for collecting and using student data: What is the end goal? Similarly, these groups must all agree with the means of implementing analytics and AI, and that they are fair and ethical.

The education leaders involved in Project Constellation describe the purpose that drove their decisions to begin working with data in a systematic way.

*Whatever we do affects millions of people—not only those that are in school now, but parents and future generations. When we take any action, we have to understand that there is a great responsibility for tomorrow. Education is a system where you see changes very slowly, but when you plant a seed, it’s very difficult to roll it back. Before making any action, it’s important to understand what values you follow and why you do it.*

_Vüsal Khanlarov_
Head of Bureau on ICT for Education,
Ministry of Education of the Republic of Azerbaijan

*A lot of our changes were driven by equity. You work so hard in a district with high poverty to make sure all student groups understand math and language arts, that you don’t realise that what employers hire people on is their collaboration skills. Can you solve tomorrow’s problem and then the next day’s problem differently? Creativity, collaboration, project-based learning: the danger is students who need these skills the most to compete in the workforce and to succeed in college are the high poverty students. But often the equity issue drives us away from the quality education that better prepares students for the #as-yet-imagined-futures.*

_Philip Neufeld_
Executive Officer, Information Technology,
Fresno Unified School District
What these leaders and their systems have in common is a shared purpose to use data to ensure that all their students are given the best possible learning opportunities to prepare them for their futures. Both their commitment to this learning analytics journey, and the specific ways they will apply learning analytics, are guided by this shared purpose.

In each system, the first questions that must be asked are what kinds of decisions analytics can support and what data elements would be useful to improve learning in specific contexts. These questions inform the end-to-end process of learning analytics and define how data should be used.

**Learning analytics to accelerate improvement of remote learning**

Remote learning has been in existence since the days of correspondence courses in which coursework and assessments were delivered by post. However, it was not until over 150 countries closed their schools during the height of the COVID-19 pandemic that remote learning shot into the spotlight as the only way in which children worldwide could continue their education. As school systems begin to reopen to in-person learning again, education leaders at all levels are seeking to take advantage of the technology investments, new practices, key learnings, and data made available from this period.

Systems of remote or hybrid learning provide advantages in terms of adding flexibility to teaching and allowing learning to occur anytime and from anywhere. But the learnings from this period can also help countries tackle some of the longstanding issues with traditional classroom education. For example, in Azerbaijan, the system of virtual schooling established during the pandemic has provided a level of transparency for all stakeholders into the usual “black box” of education.

> *This virtual school allowed Azerbaijan to establish a data-centric environment where both the school teachers, the school principals, we as a Ministry or even a parent at home can see in real time how his or her student is performing in their daily classes and how the teachers are teaching. It created complete transparency of the education process, where parents became part of the education system and they could see how the teachers are teaching and how their kids are performing while they are in this online learning environment.*

**Vüsal Khanlarov**  
Head of Bureau on ICT for Education,  
Ministry of Education of the Republic of Azerbaijan

These models for learning might also challenge the current education paradigm that has existed in many countries since the industrial age. Fresno Unified School District has been working for years to provide a 21st Century education to its students and modern technology resources to its schools. As all teachers and students now have access to devices that they use every day, the kind of data that has become available to decision-makers at all levels of the system has changed how they can
imagine education in their schools and how quickly they can see what is working to guide continuous improvement.

*If you walked into a classroom five years ago, you would see a student with a book in front of them. What were they learning? How was their mind and their spirit interacting with the learning? They might write some notes, they might write on a piece of paper, they might bubble-in a test. And maybe the test reflected a good or positive result, but we won’t know right away.*

*Now assessments occur online and give immediate feedback to the teacher and to the students. Students in our classrooms are actually choosing what next to do based on what the assessment tells them. When students are working in Teams and O365, teachers can see them working on the assignment in real time or interacting with peers in a breakout room and can provide live feedback. Now we have so many more opportunities and all of these digital signals to tell us where things are working, for which students, using which apps, and which teaching units are making the most impact. We are standing at a place of opportunity and responsibility. That’s what Project Constellation and our work with other countries is for: to make sure we move forward wisely, ethically and in a way that helps move every student and teacher forward.*

Philip Neufeld  
Executive Officer, Information Technology,  
Fresno Unified School District

**Understanding and applying the principles of responsible AI in learning analytics**

Given this rapid adoption of digital learning tools and platforms, and the new level of data it enables, educators at all levels of education systems need to make increasingly complex decisions regarding when to collect student data, which data sources to use and how to use insights about students. So, as part of Project Constellation and Open Education Analytics, we are developing a process to apply principles of responsible data and AI through our shared work.

There is always a risk that data could be used in ways for which students, parents, or other stakeholders have not agreed. Take, for example, the issue of determining the scope of student data that should be collected and used by governments. Educational data can extend into the home or private lives of students when they use their mobile or personal devices for schoolwork. At what point does use of these data violate a student’s privacy? Are students and their parents aware of how personal data might be used to make educational decisions that could impact their future? At what level and for what purposes should student data be anonymised to enable research and model-building, and who should have access to that anonymous data?
These are challenging questions. The academic literature indicates that one of the most significant risks associated with the use of learning analytics in education is the lack of universal set of principles to govern their use.\textsuperscript{xiv} A recent review found 84 unique sets of ethical principles for AI created by government, multi-national organisations, research institutions, national committees or institutes and industry in countries around the world.\textsuperscript{xv} Although these guidelines share many common elements, there is great difference in how these policies are interpreted, which are the most important, and how they should be implemented.\textsuperscript{xvi} Furthermore, there is no one body holding users accountable for upholding these principles.\textsuperscript{xvii}

Microsoft is one of the aforementioned organisations that has created and published Responsible AI Principles, for its own product development. These principles are well-aligned with those published by UNESCO\textsuperscript{xviii} and the OECD\textsuperscript{xix}. (See this paper’s Appendix for a mapping of the Microsoft, UNESCO, and OECD principles of responsible AI.) Recognising the importance of including ethics in decision-making from the very beginning, Microsoft embeds these principles in all stages of product development that involves AI.

<table>
<thead>
<tr>
<th><strong>Microsoft Responsible AI Principles</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fairness</strong></td>
</tr>
<tr>
<td>Transparency</td>
</tr>
</tbody>
</table>

Project Constellation is applying these principles of Responsible AI to develop a stronger and shared understanding of how to use these principles in education.

- **Fairness**: AI systems should treat everyone in a fair and balanced manner and not affect similarly situated groups of people in different ways. Human decision makers are susceptible to many forms of prejudice and bias, such as those rooted in gender and racial stereotypes. To ensure AI models are trained in a way that does not embed or re-enforce those biases, models must be tested for fairness. Microsoft has developed an open-source toolkit to support this called Fairlearn.\textsuperscript{1} This toolkit is incorporated in the responsible AI process that Project Constellation is developing.

- **Reliability and Safety**: Systems should operate reliably and safely when they function in the world. AI systems must be designed with a view to the potential benefits and risks to different

---

\textsuperscript{1} Fairlearn: A toolkit for assessing and improving fairness in AI - Microsoft Research
stakeholders and undergo rigorous testing to ensure they respond safely to unanticipated situations and do not evolve in ways that are inconsistent with the original shared purpose.

- **Transparency:** Transparency requires visibility into all levels of decision-making and design of an AI system. Designers should clearly document their goals, definitions, and design choices, and any assumptions they have made. Those who build and use AI systems should be forthcoming about when, why, and how they choose to build and deploy them, as well as their data and systems’ limitations. Information should be readily available on the quality of the predictions and recommendations the AI system makes. Transparency also encompasses intelligibility, which means that people (in this case, educators, parents, students, etc.) should be able to understand, monitor, and respond to the technical behaviour or recommendations of AI systems.

- **Privacy and Security:** Private or personal data should not be collected or incorporated in analytics or AI products for education unless all groups have agreed this data is necessary to achieve the shared purpose of a specific analytics or AI project. Additionally, the people providing the data need to have given explicit permission for the data to be used for this purpose (and understand the value that will be generated as a result of data sharing). Finally, the security of that data must be protected, and guidelines developed for those who can access which data, the level of anonymization needed for each specific analytics purpose.

- **Inclusiveness:** The datasets used in learning analytics and AI determine the insights and predictions produced. If those datasets do not represent the whole population of learners, if the data quality is poor, or if certain types of data are not included in the models, it will decrease the accuracy, validity, and inclusiveness of the insights. Similarly, if the way the insights are acted upon by the system do not include all groups (e.g., students with disabilities), it can reinforce exclusion from learning opportunities.

- **Accountability:** Accountability requires that people who develop and deploy AI systems be held responsible for how they operate. AI systems should never be left to operate unchecked, irrespective of the degree to which they may be capable of acting autonomously. This is what is meant by the phrase “humans in the loop.”

Microsoft has developed an understanding of how to apply these principles to the design and use of large data systems through its own product developments. Another paper in this series will describe how these principles are being applied in education through the process and technologies being developed through Project Constellation and Open Education Analytics.
Next Steps in the Journey

The education systems participating in Project Constellation describe a fruitful partnership with Microsoft and with each other that has allowed them to learn and innovate in a way that they would not have been able to if they were on this journey alone.

_It has been very helpful to work with other countries during this process. Countries may have different data and student information systems, but the learning is the same. In terms of being able to innovate, it has been very important to have the possibility to have a dialogue with others on the same journey. It is very difficult to find this kind of community within one country—to be one rebel without others to support you._

_Pasi Silander_
Head of ICT Development Programs, AI & Data Analytics, City of Helsinki, Education Division

The Project Constellation collaboration is in its early stages. There are not yet examples of full implementations of AI and their impact on student learning. Given the urgency of understanding all we can about how learning processes are changing during current global transitions in learning, the next paper in this series will present initial findings from shared learning analytics and AI models developed to address those topics. A third paper in the series will provide specific guidance on applying the Principles of Responsible AI in learning analytics.

Education system leaders are the guardians of their students and teachers’ data. Learning how to responsibly govern and use that data and the tools of learning analytics and AI for the continuous improvement of teaching and learning in education will be an ongoing journey. Project Constellation and Microsoft Education are committed to supporting this journey. We hope to learn from and with other education systems on this journey in the coming months and years.

For questions or expressions of interest, contact: openeduanalytics@microsoft.com.
### Appendix: Mapping Responsible AI Principles Across Frameworks

<table>
<thead>
<tr>
<th>Microsoft Responsible AI principles</th>
<th>OECD AI Principles</th>
<th>UNESCO Ethics of AI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fairness</strong></td>
<td>AI systems should be designed in a way that respects the rule of law, human rights, democratic values and diversity, and they should include appropriate safeguards – for example, enabling human intervention where necessary – to ensure a fair and just society.</td>
<td>Fairness and non-discrimination: AI actors should promote social justice, by respecting fairness.</td>
</tr>
<tr>
<td><strong>Inclusiveness</strong></td>
<td>AI should benefit people and the planet by driving inclusive growth, sustainable development and well-being.</td>
<td>Multi-stakeholder and adaptive governance and collaboration: Participation of different stakeholders throughout the AI system life cycle is necessary for inclusive AI governance, sharing of benefits of AI, and fair technological advancement and its contribution to development goals.</td>
</tr>
<tr>
<td><strong>Reliability &amp; Safety</strong></td>
<td>AI systems must function in a robust, secure and safe way throughout their life cycles and potential risks should be continually assessed and managed.</td>
<td>Unwanted harms (safety risks) and vulnerabilities to attacks (security risks) should be avoided throughout the life cycle of AI systems to ensure human and environmental and ecosystem safety and security.</td>
</tr>
<tr>
<td><strong>Transparency</strong></td>
<td>There should be transparency and responsible disclosure around AI systems to ensure that people understand AI-based outcomes and can challenge them.</td>
<td>Transparency and Explainability: The transparency of AI systems is often a crucial precondition to ensure that fundamental human rights and ethical principles are respected, protected and promoted. Explainability is closely related to transparency, as outcomes and sub-processes leading to outcomes should be understandable and traceable, appropriate to the use context.</td>
</tr>
<tr>
<td><strong>Privacy &amp; Security</strong></td>
<td>AI systems must function in a robust, secure and safe way throughout their life cycles and potential risks should be continually assessed and managed.</td>
<td>Privacy, a right essential to the protection of human dignity, human autonomy and human agency, must be respected, protected and promoted throughout the life cycle of AI systems both at the personal and collective level.</td>
</tr>
<tr>
<td><strong>Accountability</strong></td>
<td>Organisations and individuals developing, deploying or operating AI systems should be held accountable for their proper functioning in line with the above principles.</td>
<td>Human oversight and determination: It must always be possible to attribute ethical and legal responsibility for any stage of the life cycle of AI systems to physical persons or to existing legal entities.</td>
</tr>
</tbody>
</table>
Sources for the definitions of Learning Analytics and AI in the box are, in order of citation:
Jobin