Working Group Report on Data for Learning

The Transformative Potential of Data for Learning

September 2023

BROADBAND COMMISSION
FOR SUSTAINABLE DEVELOPMENT

Working Group Report on Data for Learning

The Transformative Potential of Data for Learning

September 2023

Acknowledgements

This report was written collaboratively, with members of the Broadband Commission for Sustainable Development's Working Group on Data for Learning contributing their knowledge and expertise. The Working Group was chaired by Audrey Azoulay (Director-General, UNESCO) and coordinated by Borhene Chakroun, Director of the Division of Policies and Lifelong Learning Systems, UNESCO, who guided the production of this report.

The UNESCO writing team included Juliette Norrmén-Smith and Saurabh Roy, with contributions from Carmela Salzano, Patrick John Devaney and Wayne Holmes. The report would not have been possible without the guidance of the Broadband Commission Secretariat team, including Anna Polomska, Leah Mann and Julia Gorlovetskaya.

The views expressed here are not attributed to any one organization or individual, except in the case studies contributed by respective Working Group members, nor do the views necessarily reflect the position of the Broadband Commission members or their affiliated organizations. Not all Working Group members participated in every meeting that led to the content of this report. We wish to thank the Working Group's Broadband Commissioners, commissioners' focal point members and external experts for their invaluable contributions, kind reviews and helpful comments.

We are grateful to the team from Strategic Agenda, including Liyana Aini, Gabriel Polley and Lauren Harper, Darcy Bush, and Julia Spaeth and Robert Tubb, for the editing, design and layout of this report, and the accompanying job board. The production of both documents was overseen by Juliette Norrmén-Smith.

Disclaimer

Working Group members

Commissioners and focal points

- Dr Tedros Ghebreyesus, World Health Organization (represented by Mr Derrick Muneene)
- Baroness Beeban Kidron, 5Rights Foundation (represented by Ms Sonia Livingstone)
- Mr Lacina Koné, Smart Africa (represented by Ms Aretha Mare, Ms Thelma Efua Quaye, Mr Calvain Nangue and Mr Osman Issah)
- Professor Jeffrey Sachs, The Earth Institute, Columbia University (represented by Dr Radhika Iyengar)
- Professor Mercedes Aráoz, Peru
- Dr Lee Yee Cheong, Former Broadband Commissioner, International Science Technology and Innovation Centre for South-South Cooperation, UNESCO
- Mr Amir Dossal, Global Partnerships Forum
- Mr Bocar Ba, SAMENA Telecommunications Council
- Mr Erik Ekudden, Ericsson (represented by Ms Zohra Yermeche)
- Mr Achim Steiner, United Nations Development Programme (represented by Mr Gayan Peiris)
- Dr Rumman Chowdhury, Parity Responsible Innovation Fund
- Dr Qu Dongyu, Food and Agriculture Organization of the United Nations (represented by Mr Maximo Torero, Mr José Rosero Moncayo and Mr Henry Burgsteden)
- Mr Mauricio Ramos, Millicom
- Mr Piotr Dmochowski-Lipski, EUTELSAT IGO
- Ms Sima Sami Bahous, UN Women

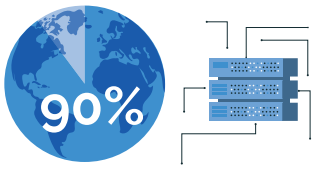
External experts

- Ms Maria Florencia Ripani, Fundación Ceibal
- Mr Dale Allen, DXtera Institute/EdSAFE AI Alliance
- Mr Alexandre Barbosa, Cetic.br, UNESCO Regional Centre for Studies on the Development of the Information Society
- Mr Fabio Senne, Cetic.br, UNESCO Regional Centre for Studies on the Development of the Information Society
- Mr Werner Westermann Juárez, Library of National Congress of Chile
- Ms Carmela Salzano, Global Partnership for Education
- Mr Amit Prakash, Centre for IT and Public Policy, International Institute of Information Technology Bangalore
- Mr Wayne Holmes, Institute of Education, University College London
- Mr Doug Casey, Connecticut Commission for Educational Technology
- Ms Karine Sonigo, International Labour Organization
- Ms Beth Havinga, European EdTech Alliance
- Mr Ashok Goel, National AI Institute for Adult Learning and Online Education
- Dr Joanna Rubinstein, former Broadband Commissioner and Council Member of the UNSDSN
- Mr Shiloh Naiken, National Education Collaboration Trust, South Africa
- Ms Tracey Burns, National Center on Education and the Economy
- Ms Pascale Raulin-Serrier, Commission Nationale de l'Informatique et des Libertés
- Mr Xavier Delporte, Commission Nationale de l'Informatique et des Libertés
- Mr Jun Yu, National University of Singapore
- Ms Velislava Hillman, London School of Economics
- Mr Stéphan Vincent-Lancrin, OECD
- Ms Carol O'Donnell, Smithsonian Science Education Center
- Ms Kati Tiainen, Microsoft
- Ms Ellen J. Helsper, Department of Media and Communications, London School of Economics
- Ms Lidija Kralj, European Union
- Mr Shitanshu Mishra, UNESCO Mahatma Gandhi Institute of Education for Peace and Sustainable Development

Executive summary

Data for learning: A double-edged sword

While education data has long been collected, recent advances in data-fuelled digital technologies and global telecommunications networks have significantly amplified the volume and complexity of learning data flows. Globally, data production is growing exponentially, with 90 per cent of the world's digital data generated in the last two years alone. Recent estimates predict that the world will generate **175 trillion** gigabytes of data per year by 2025, which is almost 88 times greater than the amount produced in 2010.



of the world's data was generated in the last two years alone
(Statista, 2022)

Despite this global upward trend, data in the education sector is only valuable if it is collected and shared in a timely, transparent and trustworthy manner, and used for the specific purpose of improving learning, teaching, administration or strategic management.

Today, education and training systems around the world are struggling to deploy their agency to steer the data revolution on their own terms, towards targeting persistent education challenges and strengthening system readiness for a rapidly changing world.

This is due, in part, to deep digital divides within and across countries. Globally, **one in three people** do not use the Internet, blocked by lack of access, affordable data, or quality connection. Despite significant international efforts to connect every school to the internet, as exemplified by the progress made by the UNICEF-ITU Giga initiative, universal school connectivity remains limited, as **25 per cent** of primary schools worldwide do not have electricity. As a result, digital education data do not flow from many rural and developing areas, creating stark inequalities in the ability to use learning data to generate insights to improve the quality of teaching and learning and include the most marginalized.

These systemic obstacles are significant. With an estimated **773 million** illiterate adults and **244 million** young people out school – in addition to the hundreds of millions of people in the workforce who require retraining – all levels of the education workforce need data competencies combined with access to quality data to: make informed decisions on how best to manage policies, budgets, resources and classrooms; adapt to the digital transformation occurring in all sectors; and develop innovative solutions to counter the global learning crisis.

There is a clear need to build the capacities of the education and training workforce to have the agency to steer the data revolution in the service of global educational equity, quality and inclusion. Supporting the growth of the sector's holistic data literacies – beyond technical expertise to include competences related to governance, regulation, compliance and social impact – requires strengthening alliances with the private sector, civil society and research institutions.

> **"If harnessed properly, the digital revolution could be one of the most powerful tools for ensuring quality education for all and transforming the way teachers teach and learners learn. But if not, it could exacerbate inequalities and undermine learning outcomes."**
>
> *United Nations Secretary-General António Guterres at the Transforming Education Summit, 2022*

In September 2022, the United Nations convened the **Transforming Education Summit**, reigniting a collective commitment to lifelong learning as a

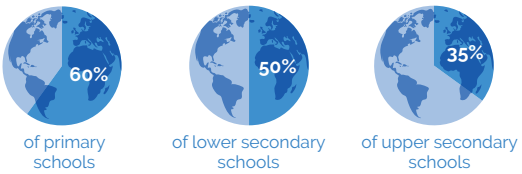
pre-eminent public good, with the urgent goal of transforming education to be relevant and responsive in the digital era. The Secretary-General of the United Nations cautioned that the digital revolution could exacerbate inequalities and undermine learning outcomes if the international community did not urgently come together to overcome the digital divide and reinforce the capacity of education and training systems to steer the process and promote its agency including through expertise, capabilities and institutions, as well as high-quality contextualized digital learning content, backed by strong, sovereign and secure data systems.

Designing, implementing, governing and monitoring such information systems is no simple task in the complex education sector, the many asymmetries of which impede synergized approaches to data for learning.

Asymmetry of information: Swift and sweeping social changes vs slow and siloed data flows

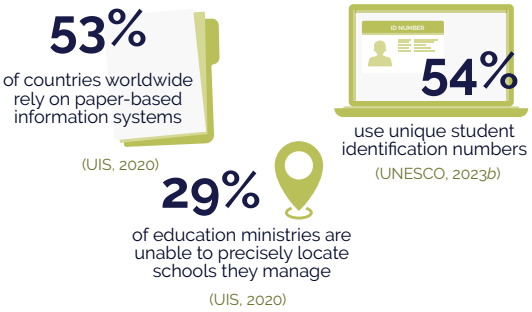
Education systems face growing pressures to rapidly respond to the many transitions unfolding around the world, from demographic and labour-market shifts, to numerous digital revolutions, to climate change. However, the sector does not always have – or use – the available information to anticipate the impact of these changes. As a result, education management often becomes a reactive rather than proactive process, unprepared to mitigate crises.

Much of the world does not have access to timely, quality and complete education data. Globally, **60 per cent** of primary schools, **50 per cent** of lower secondary schools and **35 per cent** of upper secondary schools are not connected to the Internet (UNESCO, 2023*b*). These data make clear that the effective and equitable use of digital data for learning is deeply intertwined with the digital divide. Disconnected learners, as well as marginalized learners and women and girls, risk being under-represented in education sector datasets. The invisibility of certain learners and disproportionate representation of others in these datasets may result in the further marginalization of disconnected communities whose needs remain unseen by policy-makers.



While the right to education is increasingly recognized as lifelong, beginning at birth and continuing through life, data collected in different learning settings – from schools, to workplaces, to community centres – are often disconnected, non-interoperable or lacking comparable standards. As a result, an individual's accomplishments or competencies may not be recognized or transferable, thus impeding their access to the labour market or further learning opportunities. This gets further exacerbated in the case of learners who are displaced due to forced or unforced migrations, and often lack physical records.

If information flows were better coordinated and connected both vertically across all levels of the education sector, and horizontally with the data collected by other sectors, such as health, finance, communication and labour-market data, education leaders and institutions could strengthen the resilience and resonance of their policies and practices. National, regional and global information observatories, such as those on labour-market changes, population shifts, climate risks, evolving technologies or Sustainable Development Goal (SDG) benchmarking, can be valuable analytic and comparative tools for the education sector to introduce into regular policy-making and planning cycles.



Asymmetry of skills: Big demands for "data-*driven*" vs little literacy for "data-*informed*"

There is growing pressure to use data in decision-making at every level of the education system. Therefore, every actor in the education system should have the necessary data-related competences to ensure effectiveness, accountability, compliance, privacy and security in the use of education data to improve quality and equity in all learning settings. While *data-driven* decisions tend to be based on a direct surface reading of data, *data-informed* approaches are inferred from a contextualized and critical interpretation of education data – one that balances the data with human insights, and shared principles of ethical and responsible use.

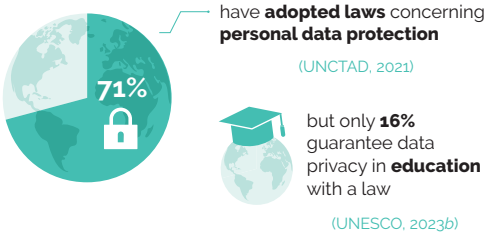
Many are adopting skills frameworks developed by the private sector that may prioritize technical skills over competences around data governance, ethics and social impact.

(UNESCO, 2023*b*)

To date, only **54 per cent** of countries have established digital skills standards, and many are adopting skills frameworks developed by the private sector that may prioritize technical skills that are valuable in commercial settings (UNESCO, 2023*b*). However, digital data literacy is not limited to technical or software expertise, such as a strong understanding of data analytics tools, cybersecurity practices, and emerging technologies such as artificial intelligence (AI) and virtual reality (VR). It is also essential to be well-versed in the social implications of data use in education, and in particular, the challenges it raises regarding inclusion, equity, ethics, ownership, agency of teachers and learners, and environmental and financial sustainability. Critically assessing the benefits and risks of data use in every learning experience is a key component of data literacy, which must be grounded in an understanding of what data represents, what it does not represent, and, indeed, what it may misrepresent.

Both critical data skills and technical digital competences are lacking around the world, posing significant barriers to the safe and effective use of data in education. Given algorithmic technologies fuelled by large amounts of data, such as applications of AI, are increasingly integrated into education systems, it is necessary that all education stakeholders are able to communicate data insights, assess data quality, grasp the main principles of data governance and ownership, and understand the impact of data use on people and human rights. Investments in capacity-building would reinforce the agency of education and training institutions to direct, design and drive the data revolution towards improving learning opportunities for all.

Asymmetry of sovereignty: Local legal frameworks vs global data flows



Data in education are not static. They are collected, processed, transferred, stored, combined, separated, archived and/or destroyed, often in different settings. These settings may not be the same classroom, school district, country or even continent. A growing concentration of data is in the hands of a few large corporations and countries located in the global North. Without adequate legislation or public awareness, the current imbalance in data power could restrict user autonomy over their data and compromise national data sovereignty and security. To avoid undermining education as a universal human right, it is crucial to assess how their support can benefit resource-scarce education systems.

> **"Despite the desire to make education a global common good, the role of commercial and private interests in education continues to grow, with all the ambiguities that entails: to date, only one in seven countries legally guarantees the privacy of educational data."**
>
> *Audrey Azoulay, Director-General of UNESCO*

Under the right conditions, cross-border data flows within a lifelong learning perspective could expand the scope of educational opportunities and foster more culturally diverse and inclusive learning environments. For example, students and educators from different backgrounds could connect and access a broad range of educational resources from various countries, including online courses, digital libraries and learning materials. Global standards for safe and secure education data sharing could also improve learner and worker international mobility, paving the way for improved mechanisms for the cross-border recognition of learning outcomes.

> **Governments should adopt and implement legislation, standards and agreed good practices to protect learners' and teachers' human rights, well-being and online safety, taking into account screen and connection time, privacy, and data protection; to ensure that data generated in the course of digital learning and beyond are analysed only as a public good; to prevent student and teacher surveillance; to guard against commercial advertising in educational settings; and to regulate the ethical use of artificial intelligence in education.**
>
> *(UNESCO, 2023b)*

For these potential benefits to be unlocked, multilateral cooperation and solidarity are necessary to navigate complexities such as intellectual property legislation, licencing and cybersecurity. Despite the global nature of data flows, minimal international cooperation on data governance and regulation impedes the realization of data for learning as a common, public good. Cross-border, multilateral normative instruments can establish ethical principles for cooperation between public and private institutions, for these partnerships can spur innovation and bolster operational capacities to develop and deliver targeted management and learning tools designed to improve education ecosystems. However, given the role education data may play in business plans, it is the international community's collective responsibility to ensure that access to information on learning and education more broadly is not only reserved for those who can pay for it, and that the knowledge commons remains a shared, community-owned resource.

In the face of these asymmetries, three essential questions guide this report on the transformative potential of data for learning:

1. **How can we maximize the benefits and minimize the risks of data use in education?**
2. **How can we ensure data is used to improve quality and equity in all learning settings?**
3. **How can we work together as an international community to take collective action to close the digital data divide, while protecting our human rights?**

On the practical level, cases of data for learning use are not easily classified as an opportunity or a risk, as their impacts may not fall on only one side of a clean-cut binary between positive and negative. After two years of debate and discussion, the WGD4L takes the position in this final report that data for learning is a **double-edged sword** that must be wielded with great care, purpose and intention. On the one hand, the education workforce needs accurate, timely and complete data to inform relevant and inclusive education practices and policies, but on the other, the use of education data cannot encroach on the rights of learners or deepen global learning inequalities. These key tensions are discussed in the report and summarized in the following diagram.

INTEGRATED DATA SYSTEMS

On one hand, ensuring that learning data is integrated across periods and sources can support learner mobility, smooth the transition from school to work, and improve government services responding to children vulnerability to harm, such as asylum-seeking children or undocumented youth. Data can play a crucial role in identifying schools in need of immediate supervisory support, for both operational, and teaching and learning purposes.

INTEROPERABLE DATA SYSTEMS

One of the major trends in education data infrastructures is to house information management systems on the cloud, with interoperability between platforms geared towards learning (LMS) and those with a broader scope geared towards management (EMIS). Enhancing cloud service interoperability can increase the potential for context-sensitive school management and monitoring of policy impact at the school level.

DATA-INFORMED PEDAGOGY

Teachers and school leaders may benefit from sophisticated data systems that can alert them when a learner's data displays early warning signs for course failure or drop-out. They can then combine predictive analytics with their own qualitative and contextualized insights to engage and support students or their families, reducing the likelihood of early school leaving. Learning management systems (LMS) can support teachers to identify individual and class trends in performance to differentiate instruction and prepare targeted lessons.

REGULATION

Strong legal frameworks are needed to protect learner data, regulate the involvement of commercial providers and procurement procedures, and prevent algorithmic bias, data surveillance and or misuse of education data. Strong regulation can ensure transparency, explainability and accountability in every education data use case.

MULTISTAKEHOLDER PARTNERSHIPS

Bold alliances between government and industry can accelerate the capacity development and growth of data and cybersecurity expertise in education and training systems. Partnerships between public and private sector stakeholders can lead to cost-effective, national-scale sustainable systems for the safe and effective coordination of data for learning.

DATA FOR ALL LEARNERS

The global community must accelerate efforts to bridge the digital and data divide that exists both within and across countries to ensure the benefits of data-reliant technologies and responsive data information management systems can reach all learners everywhere. Marginalized learners must be represented in datasets to enable inclusive digital tools and data-informed policies.

SURVEILLANCE OR BIAS

On the other hand, such integrated systems that provide a seamless transfer of comparable records between institutions could spur unnecessary surveillance mechanisms by tracking years of granular student data that may make negative data points (behavioural or academic) difficult to recover from, thereby influencing negative institutional and self-perception of students' success in learning systems.

COMMERCIAL CONTROL

However, this architecture also favours the market by lowering the cost of entry into existing digital infrastructures, in turn, can lead to vendor-controlled pricing of data services. It also greatly expands the amount of data that can be collected through the streamlining of data pipelines and single sign-on capabilities that can collect user data regardless of location (see, for example, Gulson *et al.*, 2022).

HYPER-DATAFICATION

Collecting data for the sake of it can create high-stress, high-stakes school environments. Limited training on data processes, combined with the pressures to continuously collect for accountability's sake, can result in teacher frustration and a decline in the attractiveness of the teaching profession (UNESCO, 2017). Technology in education should put learners and teachers at the centre to prevent it becoming a burden or an impediment to teachers' ability to teach (UNESCO, 2023*b*, p. 21).

INNOVATION

Over-complicated and opaque legal and regulatory requirements for the use of education data can prevent local companies from developing relevant digital learning solutions, preventing the growth of local education technology ecosystems that create contextualized data-fuelled digital tools for education.

CONTROL AND SUSTAINABILITY

Governments may yield their agency in steering the safe and sustainable use of education data use if essential data skills are outsourced to private providers for data harmonization, compilation, analysis of raw data, development of application programming interfaces, creation of dashboards or custom reports, and cybersecurity provisions, among others. Less than one third of countries have a sustainability clause in their procurement law.

FINANCIAL AND ENVIRONMENTAL COSTS

Data generation and analysis comes with high financial and environmental costs. Estimates indicate that affordable data is the most costly component of financing universal digital learning, with a price of $498 billion for the period 2021–2030 (UNICEF, 2021). Data centres account for a large percentage of electricity demand and produce a heavy carbon footprint.

These tensions are not easily resolved, meaning that it is of paramount importance that all education decision-makers design and review policies, practices and architectures that maximize the potential and minimize the risks of data use in education. Unleashing the potential of data use in education hangs in a delicate balance, not a binary, between benefits and risks, data-driven and data-informed, innovation and regulation, and surveillance and invisibility.

The Broadband Commission convened a working group chaired by UNESCO and composed of commissioners and experts in January 2022 to discuss the double-edged nature of data for learning along three thematic areas: (1) **infrastructure** and **architecture** of education data ecosystems, (2) data **skills** and **competence** frameworks for life and work, and (3) **governance**, regarding ethics, national sovereignty and cross-border data flows. These areas structured the work of the group and were adopted as strands of dialogue and reflection. At each monthly meeting, working group members, as well as visiting experts and institutions, were invited to present thematic cases to enrich the discussions. To build its analysis, this report explores some of these cases for how governments and industry are handling data-related issues through their domestic education policies, and bilateral or multilateral arrangements.

The culmination of this analysis is the report's **five recommendations** that the Broadband Commission WGD4L presents to policy-makers and stakeholders engaged in the education data landscape.

Overall, this report argues for the need for multisectoral, multilateral and multi-stakeholder cooperation to safely unleash the potential of data for learning to drive the improvement of education's persistent problems. Ethical and purposeful uses of data play a key role in improving school and system management, supporting inclusive and innovative learning and teaching methods, and ensuring the equitable financing of education by illuminating areas of systemic need for investment.

Five recommendations:

Develop and implement a **whole-of-government and whole-of-ecosystem vision and strategy** on the use of data for learning, grounded in a rigorous understanding of the potential opportunities, benefits, limitations and risks.

Establish a **sustainable financing strategy** for data for learning, grounded in national financial resources, that benefits learners, promotes data in education as a public good, minimizes environmental impact, and is sustained by strategic multi-stakeholder partnerships.

Strengthen critical **data literacy and skills** at all levels of the education ecosystem to facilitate improved regulation and inclusive innovation through effective implementation and monitoring of education data policies and practices.

Prioritize uses of education data that **target systemic obstacles** to universal access to equitable and inclusive quality teaching and learning, by improving the efficiency and effectiveness of education management, administration, planning and financing.

Harness **multilateralism, solidarity** and **international cooperation** to develop international standards and norms over education data regulation, literacy, cybersecurity, governance and ethics, bridge the digital divide, nurture local data capacities, and promote free open-source software to support the development and implementation of safer and more targeted tools for the education sector.

The data on education data

97 countries lack foundational learning data to know which students can read basic text by age ten

(UIS, 2020)

US$498 billion would be needed to make data usage affordable in learning activities for the coming decade

(Yao et al., 2021; UNESCO, 2023*b*)

In **44%** of low- and middle-income countries with pricing data, the **median cost** of 1GB of data **exceeded 2% GDP per capita**

(UNESCO, 2023*b*)

In 2022 the education sector accounted for...

5% of all ransomware attacks

30% of security breaches

(UNESCO, 2023*b*)

Globally, there is no **foundational literacy data** for **18%** **of the school-age population**

In Sub-Saharan Africa, this number grows to over **50%**

(UIS, 2021)

60% of primary schools

50% of lower secondary schools

35% of upper secondary schools

are not connected to the Internet

(UNESCO, 2023*b*)

71% of countries in the world have **adopted laws** concerning **personal data protection**

(UNCTAD, 2021)

but only **16%** guarantee data privacy in **education** with a law

(UNESCO, 2023*b*)

53% of countries worldwide rely on paper-based information systems

(UIS, 2020)

54% use unique student identification numbers

(UNESCO, 2023*b*)

29% of education ministries are unable to precisely locate schools they manage

(UIS, 2020)

89% **of 164 recommended learning platforms or products** during COVID-19 were found to have tracking technologies that could or did monitor children and harvest personal data on children to send to third-party companies

(Human Rights Watch, 2022)

39 out of 42 governments that provided online education to children during the COVID-19 pandemic used digital technology in ways that violated children's rights

(Human Rights Watch, 2022)

Only 54% of countries have mapped out digital skills standards

Many are adopting skills frameworks developed by the private sector that may prioritize technical skills over competences around data governance, ethics and social impact.

technical skills | data governance, ethics, social impact

(UNESCO, 2023*b*)