

**DIRECTORATE FOR EDUCATION AND SKILLS
CENTRE FOR EDUCATIONAL RESEARCH AND INNOVATION (CERI) GOVERNING BOARD****LONGITUDINAL DATA SYSTEMS IN EDUCATION****WORKSHOP SUMMARY REPORT****"FOSTERING INNOVATION AND IMPROVEMENT IN EDUCATION : THE ROLE OF
LONGITUDINAL INFORMATION SYSTEMS"****30 June - 1 July 2014, New York City, United States**

In June 2014, the OECD Centre for Educational Research and Innovation (CERI) organised an international workshop to discuss the role of next-generation longitudinal information systems in promoting educational improvement and innovation in the era of "big data". The meeting built on the dialogue started at the 2010 OECD-SSRC-Stupski Workshop and took stock of developments in recent years. It was attended by more than 30 policy makers, data use experts and system managers from OECD Members and Non-Members.

The objective of the workshop was to foster discussion about the latest developments in the design of longitudinal education data systems, the conditions under which they are most effective and the main challenges associated with their use.

This report summarises the key points of the presentations and discussions.

Stéphan Vincent-Lancrin (OECD), EDU/IMEP, stephan.vincent-lancrin@oecd.org
Carlos González-Sancho (OECD), EDU/IMEP, carlos.gonzalez-sancho@oecd.org

JT03372415

Complete document available on OLIS in its original format

This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

TABLE OF CONTENTS

FOSTERING INNOVATION AND IMPROVEMENT IN EDUCATION: THE CONTRIBUTION OF LONGITUDINAL INFORMATION SYSTEMS	5
Introduction.....	5
Day 1: Mapping technical features, current uses and opportunities for development	5
Session 1: The state of longitudinal information systems in education today.....	6
Session 2: Information systems in higher education and links with the labour market	9
Session 3: Roundtable next-generation systems. Promising features in the R&D pipeline	11
Session 4: Setting up a new system: policy mandates and practical considerations	12
Session 5: Discussion and policy recommendations	13
Day 2: Using next-generation information systems for continuous improvement in education.....	14
Session 6: Supporting educational research	14
Session 7: Instructional practices and personalised learning	16
Session 8: Innovative approaches to school and system improvement	17
Session 9: Conditions for effective data use	18
Session 10: Next steps toward solutions. Breakout sessions and final plenary discussion	19
ANNEX: AGENDA AND LIST OF PARTICIPANTS	21
Agenda	21
List of participants	26

Boxes

Box 1. Highlights of the discussions.....	3
---	---

Box 1. Highlights of the discussions

Longitudinal information systems are supporting an increasing number of strategies for educational improvement and innovation

- Information systems are being developed with a growing diversity of goals and approaches to data use. While policy planning and accountability were the key drivers over the last decade, systems are also increasingly used to support educational research and more personalised teaching and learning practices. They could also be used for a variety of other goals and become “expert systems” for all practitioners in the system.
- As education enters the era of “big data”, opportunities for next-generation systems to act as catalysts of improvement and innovation increase, but so do challenges associated with data management, data ownership, and privacy protection.

Effective data use remains a challenge despite progress in system infrastructure

- Many countries or jurisdictions have equipped themselves with sophisticated data systems, but investments in technology and data collection have not generally been accompanied with successful strategies to promote effective data use, especially at the school and classroom levels.
- Effectively deploying a data system requires a common understanding of its multiple purpose and functions. Consultation with stakeholders is needed to clarify the key questions that the system will or could help address in the short and in the medium run.

Building partnerships and trust for data sharing and data use is crucial

- Data sharing partnerships expand possibilities to use data for educational improvement. These partnerships depend crucially on establishing institutional trust and collaboration among participating agencies and institutions. The development of information systems provides a new opportunity to set up formal procedures, share data across institutions and make access to data less dependent on personal relationships.
- Policy makers can promote changes in regulatory frameworks so that a greater integration of data resources helps address key policy questions. A critical step is the creation of individual unique identifiers allowing data matching and longitudinal tracking across several longitudinal data systems. In order to secure stakeholder buy-in, the use of information systems cannot be predicated solely on high-stakes accountability and reporting requirements. Practitioners need support to use data for improvement purposes, while policy-makers and chief data officers need to carefully examine the different drivers of data use by practitioners.
- Education data sharing and use call for sustainable data governance structures that regulate ownership and accountability and ensure continuity through policy shifts. Documenting procedures for data access and changes to the system is critical to facilitate knowledge transfer.

Information systems offer promising avenues for greater personalisation of teaching and learning

- Data mining and learning analytics hold great promise to improve the personalisation of instruction and learning. These techniques can combine data from formative assessment and informal learning activities and interactions to identify individual skills gaps and target content and instructional delivery. They can also give learners a greater autonomy in designing their learning experiences in education.
- Learning analytics applications can be used to derive insights from the vast amount of data maintained by longitudinal information systems, for instance by mining historical data to support benchmarking and statistical inference. Visualisation tools providing detailed feedback to learners and educators are other important features of these systems.

- Longitudinal information systems in combination with analytics facilitate a greater personalisation of learning and may lead to a redefinition of the role of teachers towards more immersive, active pedagogies and individualised tutoring. They can propose an array of diagnosis and enhancement tools to these systems.

Protecting privacy to enable innovative uses of education data

- Preserving privacy without hampering the use of personal data to improve educational outcomes remains a key challenge for data sharing and more specifically for research. Promising models for balancing these objectives combine strict selection criteria of trusted data users and little restrictions in terms of accessible data elements and linkage possibilities.
- Specific solutions include the establishment of a third-party authorized with holding keys and linking data from multiple agencies; the release of de-identified public use files; supervised access in research data centres; and a variety of data de-identification techniques.
- Changes in legal and governance systems have not followed the pace of developments in data generation and data use patterns and will need to address growing privacy and ethical concerns in the coming years.

New models of professional development based on using education data and information systems

- Promoting an effective use of information systems by teachers and education policy makers requires both a greater customisation of systems to different user needs and fostering data literacy independently of training for specific data retrieval and analysis tools.
- The adoption of an information system is a socio-technological process of innovation and its success depends mainly on human and organisational aspects rather than technology. Building common understandings for data use and incorporating feedback from practitioners are important to design effective systems and prevent negative disruptions in work and decision-making procedures.
- Longitudinal information systems could lead to a new culture of individual, organisational and sectorial learning in education, and thus of continuous improvement and innovation. Not only can they be equipped with social network functionalities that facilitate the establishment of virtual communities of practice, they can be used within schools to support peer learning and new forms of professional development within schools.

**FOSTERING INNOVATION AND IMPROVEMENT IN EDUCATION:
THE CONTRIBUTION OF LONGITUDINAL INFORMATION SYSTEMS**

**International workshop
OECD Centre for Educational Research and Innovation (CERI)
30 June - 1 July 2014
Barnard College, New York City, United States**

Summary Report

Introduction

1. The workshop “Fostering Innovation and Improvement in Education: the Contribution of Longitudinal Information Systems” continues the efforts of the OECD Centre for Educational Research and Innovation (CERI) to look at the role of information systems in the innovation process in education. This work explores the potential of longitudinal information systems to harness the growing volume and diversity of education data in order to promote improvement and innovation in education.

2. The aim of the international meeting was to foster discussion about the latest developments in the design of longitudinal education data systems, the conditions under which these tools are most effective, and the main challenges associated with their use.

3. The workshop built on the dialogue started at the 2010 OECD-SSRC Workshop [EDU/CERI/CD(2010)14] and took stock of recent developments in the design of information systems and the use of education data at large. It also served to present interim results from the ongoing CERI survey of longitudinal information systems in education.

4. Invited participants were a mix of policy makers, data use experts and system managers and developers from OECD countries and partner economies. Their number was limited to 40 to encourage interaction and in-depth discussion. The workshop was attended by representatives of public education agencies and private organisations from Austria, Brazil, Canada (Ontario), Estonia, Mexico, the Netherlands, Norway, Portugal, Slovenia, Spain and the United States. The full list of participants can be found in the Annex.

5. The workshop was organised by CERI within the framework of its project *Innovation Strategy for Education and Training*, and it was hosted by Barnard College, Columbia University. Stéphan Vincent-Lancrin, OECD, and Kathryn Kolbert, Athena Center for Leadership Studies, Barnard College, welcomed participants on behalf of both institutions.

Day 1: Mapping technical features, current uses and opportunities for development

6. The first day of the workshop aimed at outlining the objectives, characteristics and functionalities of existing education information systems across the OECD area and beyond. The objective was to establish where most information systems currently stand, and to highlight interesting features and uses as well as directions for future development.

Session 1: The state of longitudinal information systems in education today

7. The opening session provided an overview of the workshop aims and interim results from the CERI survey. It was followed by presentations of specific information systems in different countries.

8. Stéphan Vincent-Lancrin, OECD, introduced the key questions motivating the organisation of the workshop. The project *Innovation Strategy for Education and Training* sees longitudinal information systems as a technology tool that can be mobilised to promote innovation in the education sector in combination with other policy levers of a broader innovation ecosystem. Next-generation information systems should integrate statistical data systems and learning management functionalities with the objective to provide quick feedback and suggest solutions to education stakeholders. As education enters the era of “big data”, new opportunities and challenges emerge for using longitudinal information systems to foster educational improvement and innovation. The workshop sought to contribute to this discussion by identifying new horizons for the development and use of these systems.

9. The potential benefits of next-generation information systems should be considered across multiple domains. These benefits range from improving the speed and quality of feedback that education decision makers and practitioners receive; to supporting new forms of formative assessment and the design of more personalised teaching and learning strategies; to enabling networking platforms to mobilise practical knowledge around learning communities; to supporting the development of a stronger educational industry; to improving the efficiency of administrative operations in schools and education agencies; and to creating a better data infrastructure for educational research and the evaluation of educational innovations.

10. The analysis of how to best leverage the potential of longitudinal information systems should also address a series of critical challenges. These include how to ensure that systems are consistently fed with high quality data; how to balance evaluation and accountability purposes in order to secure buy-in from multiple stakeholders; how to prevent the emphasis on data use to induce a narrowing of educational objectives; how to create the conditions for effective data use at the school level; and how to establish models of data access that address growing privacy and security concerns about the use of personal data.

11. Carlos González-Sancho, OECD, presented interim results from the ongoing CERI survey which will yield the first international mapping of longitudinal information systems in education. As of June 2014, the survey covered 64 systems in 25 OECD and partner countries/economies. Preliminary results from 26 systems were presented at the OECD-SSRC Workshop in 2010. The survey has since collected data on another 20 systems from 16 new participating countries/economies. In addition, the survey was enriched with data on 18 longitudinal state-wide systems in the United States collected in 2013 by the Data Quality Campaign (DQC).

12. The majority of the systems in the CERI survey have been in use for less than a decade, are publicly administered and include data from national or state jurisdictions. Besides data from primary and secondary schooling, about half of the systems collect as well data from pre-primary or post-secondary/higher education, and up to a third of the systems provide comprehensive P-20 coverage. Comparisons also reveal that substantial variation exists across systems in terms of their goals, the nature of their data, their access restrictions and the functionalities that they enable.

13. Preliminary analysis suggests that the majority of current systems could improve on a number of dimensions. While most systems are able to track individual students and schools, fewer include the teacher and course identifiers required to pinpoint instructional practices where changes may be needed. In terms of data elements, most systems could make progress in incorporating richer student assessment data which to date remains largely focused on attainment and summative indicators. Systems could also extend their teacher and school level data to include professional development and evaluation records. More

generally, incorporating data from other agencies such as health care, social services or labour market data remains an area of potential development. Access restrictions could be relaxed as many systems make data available to education authorities and school leaders only, even after anonymisation procedures. In order to support improvement efforts in a timelier manner, reducing the delays in providing feedback to users would appear as a priority for most systems, as in many cases these delays amount to more than six months. Many systems could also further integrate data analysis and visualisation tools such as automated reports and customisable dashboards. Lastly, most systems could increase their ability to support changes in classroom practices by integrating links to repositories of digital resources and expert recommendation engines. Further checks on the quality of the survey responses will be needed before specific recommendations for improvement can be formulated.

14. Tadesse Haile, Ontario Ministry of Education, presented the Ontario School Information System (OnSIS), which collects data to inform decision-making related to education policies, programs and practices within the province of Ontario (Canada). OnSIS was launched in 2005-06 as part of the broader Managing Information for Student Achievement (MISA) initiative, which had the goal of building the capacity to use data at both the local (school board) and provincial levels. OnSIS collects over 100 million new data records across multiple levels and collections every year. Despite the large growth in the volume of data collected with the introduction of OnSIS, the supports provided through the MISA initiative have enabled a significant decrease in the time required to collect data from boards and an increase in the quality of data.

15. A primary function of OnSIS is to support the ministry's analytical needs and provide key indicators about policy priorities. The system enables highly granular statistical modelling and trending analysis and a rich contextualisation of student achievement patterns over time. Longitudinal indicators are constructed and may be aggregated at the provincial, school board and school levels, as well as within sub-groups. This enables improved monitoring and a better understanding of the factors relating to student attainment. Longitudinal tracking of individual students is made possible with the Ontario Education Number (OEN), a unique student identifier that provides links to data from other sources. Uses of these data include the development of indicators tracking changes in the proportion of students moving between different levels of achievement between grades, and examining trends relating to student registration in postsecondary education in Ontario. By collecting timely and quality education data, OnSIS also supports information dissemination through the production of public reports. Examples include board progress reports, a school information finder, and trend reports at the provincial level. This information provides timely and consistent evidence to inform strategic planning and decision-making.

16. Kari Steenstrup, Norwegian Directorate for Education and Training, presented the suite of systems developed by Norway to manage its education data. The Directorate runs a central database which integrates records from multiple public registers at various levels of education and enables researchers and Statistics Norway to link these to data from higher education and other agencies. The statistical unit of the Directorate conducts longitudinal analyses of these individual-level data to support education policy making across a range of areas and initiatives.

17. By legal mandate, Norway makes public all its education data in aggregated form. Two platforms have been designed for this purpose. *Skoleporten* is a web-based tool presenting selected school-level indicators and tailored reports to local education authorities and school principals to help them benchmark their performance and support local assessment of quality in education and school development. A separate system currently under construction, *Statistikkportalen* will contain a richer set of data and serve broader aims. General users will be able to customise reports by choosing variables and units of analysis for comparison and to export aggregated-level data. Eventually the two systems will become integrated with *Statistikkportalen* operating as the only data source for both portals.

18. Melissa Beard, Washington State Education Research and Data Center, presented the longitudinal data system of the state of Washington (United States). The system integrates records from multiple agencies across the state into an operational data store with over one billion records linking individuals from early childhood education through K-12, post-secondary and into the workforce and providing information on a wide range of programmes and outcomes. The system was designed to improve research capacity to address critical questions for education policy at the state and local levels through the production of feedback reports and serving as a data source for educational researchers. Washington's approach to facilitate data sharing was to establish a multi-agency data governance structure and to rely on a third party vendor for software to assist in linking identities in its data warehouse.

19. Frank D. Barnes, Charlotte-Mecklenburg Schools (CMS), presented the information system of CMS, the 17th largest school district in the United States serving 146,000 students across 189 schools. The system aims to facilitate data-informed decision making for school improvement. Besides supporting management and accountability at the district level through quarterly and annual school performance reviews, the CMS information system puts data in the hands of practitioners with a strong focus on promoting changes in practices at the ground level. Principals and teachers can use the system to examine a wide array of longitudinal student-level data including interim assessment results and attendance and incidence records that are updated daily. These highly granular indicators are the key input for designing and iterating short instructional cycles in a formative and low-stakes approach to data use at the school level.

20. Birgit Lao-Petersoo, Estonian Ministry of Education and Research, presented the Estonian Education Information System (EHIS). EHIS was established in 2005 as part of Estonia's extensive e-government infrastructure. EHIS and the rest of the state's information systems use a data exchange layer to automate data sharing in a secure Internet-based environment and are accessible to citizens through their ID digital cards. By integrating data from different education registries, EHIS allows individuals to access personal education and training records across most of the life course (adult training data is still partial and under development). The data opening web environment *HaridusSlim* (Education Eye) within EHIS provides information about education programmes and institutions across the country in order to support decision making at multiple levels, from school choice by students and families to planning and monitoring by local and state policy makers. Future plans include launching a new version of EHIS and enriching the information contained in *HaridusSlim* to enable longitudinal analysis of school effectiveness and graduates' success in the labour market.

21. Presentations from systems stimulated much discussion throughout the morning sessions. A recurrent theme was how to first identify and then balance stakeholders' needs when deciding what data to collect and what indicators and reports to produce. Systems such as Ontario and Norway carry out needs analysis through regular consultation in advisory and priority setting committees where multiple stakeholders are represented, including teachers unions and parental organisations. Collections and analytical work should target the data elements and types of comparisons that these stakeholders find most useful and avoid duplicating and storing data of little relevance or novelty to them. While consensus may exist around key indicators, accountability mandates can often pose a challenge to prioritisation exercises. These can be facilitated by consultation and pre-release review processes with stakeholders directly concerned by data sharing and reporting. Systems should also be designed with a built-in agility to customise their data and analytical outputs as stakeholder needs evolve.

22. Policy shifts were also discussed as a factor having an impact on the prioritised uses of education information systems. Building relationships of trust across institutions was deemed key to prevent that such changes disrupt the continuity needed to build data infrastructure and set in place data sharing and data use protocols. Some participants noted that using data use for accountability and evaluation purposes can set new expectations for reporting that survive policy shifts. On the other hand, changes in regulatory

frameworks can facilitate the development of information systems. Legal changes to permit data sharing across agencies and the establishment of individual unique identifiers were highlighted as decisive steps for some systems, while others also cited new legal requirements for public reporting. The importance of strong leadership in advancing evidence-based approaches in education management was also mentioned. Data-savvy policy makers are aware that a poor integration of data repositories limits the ability to answer relevant policy questions.

23. Representatives of different systems noted as well that a strong rationale for investing in the development of data systems lies in the savings brought about by improved data collection and data quality assurance procedures and in an improved capacity for policy making.

24. Another key question related to placing fair expectations on the ability of any information system to serve multiple purposes with the same effectiveness. A distinction was made between ‘strategic’ and ‘tactical’ uses of data and information systems. The former would focus on providing an aggregate-level perspective on system performance with the main aim of supporting policy planning and monitoring as most state-wide systems in the United States do, while the latter would relate primarily to informing changes in practice at the classroom and school levels. As these approaches have different implications in terms of data requirements (e.g. granularity) and system functionalities (e.g. speed of feedback, access restrictions), it was argued that data systems built for specific purposes should not be expected to be equally effective on all dimensions. It was also noted, nonetheless, that much of the infrastructure and data of any given system can sustain multiple uses, as exemplified by the case of the Charlotte-Mecklenburg Schools system.

25. The question of how much and what types of data to put in the hands of the general public was also debated as data may “live their own life” after release. One position in this debate is that making most data accessible for stakeholders constitutes a valid strategy to encourage transparency, accountability and organic uses of data. As a downside, this scenario must confront the challenges of ensuring student confidentiality (as re-identification may become possible, for instance, by triangulating sources) and of handling the politics of school and teacher evaluation. Data may also be used for purposes different than those that led to their collection. This may speak to innovative data uses but also to risks.

26. As an overarching theme, the discussion also revealed a common perception that few information systems in education are currently being exploited to their full capacity. In many cases, the lack of a clear vision and consensus among stakeholders about the purpose and potential of the system rather than technical or data limitations were highlighted as the factors explaining underutilisation.

Session 2: Information systems in higher education and links with the labour market

27. Presenters in this session were invited to discuss the specific challenges and opportunities for the use of information systems in higher education which, as compared to K-12 education, may relate to the greater autonomy of higher education institutions, the greater mobility of students across institutions, and connections with the labour market and life-long learning.

28. In Brazil, the National Institute for Educational Research (INEP) has developed several information systems to promote research and evaluation of the Brazilian education system. Nara Viera, INEP, presented Censup, the system used to collect data from around 2,400 public and private higher education institutions in Brazil. Reporting to Censup on a yearly basis is mandatory for all higher education institutions, which must provide information on their programmes of study and academic and administrative organisation. From 2009, the system began also to collect individual-level student and teacher data which can be linked to other public registries using the CPF (national identifier). Censup data are mainly used to support education policy making, from allocating federal funds to public universities to

forecasting demand for higher education across regions and fields of study. All of its data is published in anonymised form and accessible to citizens through the INEP website.

29. The state of Colorado, United States, has been collecting student-level data from its public higher education institutions since 1988 through the Student Unit Record Data System (SURDS). Daniel Domagala, Colorado Department of Education, explained how Colorado has established strong partnerships across agencies to match K-12 data to the wide array of elements in SURDS including application, enrolment, remediation, completion and financial aid student data. These longitudinal matches, going back to 2009 with a success rate in the region of 94%, allow stakeholders to analyse the factors influencing student postsecondary progress and success disaggregating data by key student demographics and running comparisons between districts and state averages. Another initiative has enabled linkages with earning data for recent college graduates from the Colorado Department of Labor and Employment, which can now be examined by institution or area or study via the College Measures portal. Linkages to data from the National Student Clearinghouse permit to track Colorado high school graduates in out-of-state higher education institutions not included in SURDS.

30. Regina Radinger, Statistics Austria, presented Unidata, the data warehouse used by the Austrian Ministry of Science to support strategic planning and monitoring in higher education policy. Unidata combines data from three different databases to cover key areas of operations of higher education institutions such as budget, enrolments and graduations and provides open access to a set of dynamic web-based reports enabling comparisons across institutions. Statistics Austria established a database that incorporates individual level identifiers and serves as a source for pathways analysis on transitions within and between different education programmes. Forecasts of the number of new entrants in higher education up to 2020 have contributed to the design of the national higher education strategy currently being developed by the Austrian government. Analyses of labour market outcomes at different time points are possible through matches with other public register data. Higher education institutions rely on these analyses to design their programmes of study.

31. The open discussion following the presentations addressed the practical difficulties of matching education and labour market data. Participants remarked that linkages would in fact be possible in most jurisdictions via existing individual unique identifiers such as social security numbers, however in most countries or states legislative changes would be needed before agencies can engage in data sharing and matching. A major concern about the use of universal identifiers is that more individual data becomes exposed to re-identification in the event of security breaches. As in the case of data sharing between education agencies, the establishment of a third party authorised with holding keys from multiple databases and linking data that are then de-identified before being made available to the contributing agencies was proposed as an efficient solution for enhancing data protection and trust. Some participants noted that these arrangements should remain open to privately developed databases which may complement public registries.

32. Another theme of discussion were the different forms that the preoccupation with teaching and learning takes in K-12 and higher education, and how these may influence uses of education data. The fact that individual educators in colleges and universities are rarely subject to the same type of public scrutiny as teachers in primary and secondary schools has put less emphasis on linking teacher data to student and course data in higher education. Accountability debates around value-added measures have thus been largely absent in higher education. Some information systems data are nonetheless being used to feed back information to secondary school about the performance of their graduates in higher education, for instance on remediation needs, and are therefore supporting improvement in instructional practices.

Session 3: Roundtable next-generation systems. Promising features in the R&D pipeline

33. This session invited entrepreneurs and businesses to present promising ideas about what next-generation systems could look like in the coming years as technological innovations in the R&D pipeline as well as new business models to overcome traditional institutional barriers around data use begin to emerge.

34. Nelson Gonzalez, Declara, discussed how information systems could support deep learning through the recommendation of personalised learning paths. Technological innovations in this area seek to activate dormant information in virtual learning and collaboration environments and to derive personal cognitive graphs from learners' dynamic interactions with content and social data. Machine learning models and advanced semantic search algorithms open new possibilities to mine data generated in both formal and informal learning settings, including web navigation, blogs, and social media. For example, Declara applied item-level analytics to conversations and blog entries in a teacher online discussion platform in Australia to derive valuable insights for the design of a new curriculum. Cognitive graphs and content discovery recommendations can also be derived from learner-generated data in other interaction contexts such as MOOCs.

35. Gerri Burton, New Learning Ventures, argued that information systems in education can support innovations similar to those achieved in urban and energy management and captured by the notions of "smart cities" and "smart grids". Talent supply management to reduce the mismatch in the formation and use of skills in fast changing economies calls for a much more intensive use of education and training data, especially in higher education and work-based learning. Advanced analytics and information systems integrating data on personal learning trajectories and labour market needs could help improve communication across education and employment institutions and thereby the design of study programmes and talent management by government and business officials.

36. Chris Lohse, Renaissance Learning, noted that educational innovation is hampered by very low levels of investment and R&D and, often, by a failure to meaningfully engage with existing innovations and insights from academic research. The emergence of open data and resources has however pushed for improvement in the provision of educational products and services and a large number of actors are proposing innovative data management solutions. A common factor to many successful initiatives has been to design systems giving voice to practitioners and recognising their professional knowledge. This often involves incorporating unconventional data elements, often of a qualitative nature, that educators would like to have in the system to support their work.

37. The ensuing conversation addressed the extent to which traditional scientific and ethical standards can be maintained in light of some of the proposed innovations as the criteria that define data sources deemed valid for research and improvement efforts begin to blur. For instance, concerns about privacy, integrity and selection effects emerge about linking and jointly analysing data collected in educational settings together with open content and directly or indirectly user-generated data via web navigation, mobile devices and the like.

38. It can be argued that the proliferation of data generation methods can help provide a richer view on personal learning dynamics and educational experiences that what is currently afforded by conventional data records. Similarly, new types of data and more frequent data generation cycles may allow a more rapid and efficient testing of educational innovations than traditional experimental approaches. However, new standards are needed to guarantee research ethics and privacy protection without hampering innovation. Progress in conciliating these agendas is for instance visible in healthcare, where greater consensus exists on the potential benefits of combining clinical records and other types of personal data. It

is also important that opportunities to exploit big data in education exist for all stakeholders and not just for large agencies and corporations.

39. The importance that new products and systems make progress in terms of data visualisation was also emphasised as crucial. Trends in design are largely towards more organic visualisation models and interfaces allowing greater customisation and choice of the data elements that users find relevant to their needs.

Session 4: Setting up a new system: policy mandates and practical considerations

40. The objective of this session was to discuss the policy mandates and challenges faced by education agencies in the process of setting up or redesigning their information systems.

41. As an introduction to the presentations of specific initiatives, Tate Gould, Anlar Inc., put forward some practical recommendations derived from the experience of state-wide systems established in the United States over the last decade. Firstly, policy questions rather than technology should guide the design of information systems, with mission statements and policy questions being formulated early on to inform IT solutions and data collections rather than vice versa. Secondly, key stakeholders should be identified and consulted throughout the process following a detailed communication calendar. Thirdly, sustainable data governance structures should be put in place through a data governance charter and formal decision protocols that establish ownership and accountability. Fourthly, agencies should extensively document their outputs and procedures to facilitate knowledge transfer, for instance by producing data dictionaries and writing down protocols for data access or changes to the system.

42. Agustín Caso-Raphael and Arturo Cervantes, Mexico's National Institute for Educational Evaluation (INEE), presented INEE's plans to develop the new information system *Sistema Integral de Resultados de las Evaluaciones* (SIRE). Mexico has recently engaged in a programme of structural reforms across multiple sectors including education. Following this reform, INEE was granted institutional and budgetary autonomy in 2013 with the mandate to coordinate the evaluation of the education system at federal and state levels. The SIRE system is being designed as a tool that brings this mandate into a business enterprise. Its development plan started in 2014 with the design of the system architecture and data model, the development and consolidation phase is planned for 2015, and the final testing and implementation are expected for 2016. SIRE will integrate geographic and administrative information as well as the inventory of all preschool, primary and secondary education schools across Mexico. It will also serve to collect longitudinal individual-level data to inform teacher, student and school evaluations, and the design of public policies and directives. Another objective is to enable future extensions of analytical functionalities and interoperability with other national and local systems. One of the major challenges at present is to establish effective data governance structures that address regulatory issues

43. Teresa Evaristo, Portuguese Ministry of Education and Science, discussed Portugal's efforts to integrate its fragmented education data infrastructure into a new centralised information system called SIGA. Portugal has been compelled by the recent economic crisis to improve its capacity to evaluate the efficiency and efficacy of its education system. The country had previously developed the MISI information system to collect data from independent school management systems on a monthly basis and feeding a separate database per school year. Despite containing very rich data at the individual and school levels, the system did not enable longitudinal linkages and suffered from poorly defined standards. For these reasons, usage was largely limited to producing yearly reports for budgetary and policy planning.

44. The new SIGA system seeks to overcome these limitations and strengthen the ability of the ministry to produce improved indicators and address pressing questions such as high grade retention and dropout rates by tracking student pathways in a more detailed and timely manner. In this respect, SIGA

will benefit from the national e-enrolment system based on electronic IDs introduced in 2010. Individual identifiers will also permit to link student, teacher and course data across school years. In addition to the new collections, the MISI central database will incorporate old cross-sectional data series, which have been matched with a success rate close to 80%. The ministry has consulted schools on the design of the new system, which will become mandatory for them to use. SIGA will aim to facilitate school management by automating tasks such as certificate issuing through the portal, as well as to support improvement efforts through early warning tools for educators. Future projects also involve the use of learning analytics to improve instructional practices. SIGA will be accessible through the *Portal des Escolas* website.

45. How to engage data providers and target audiences in the early development stages of new systems was a key question in the ensuing debate. Policies to foster this engagement may include carrying out extensive consultation about user needs during design and early implementation phases; providing valuable feedback as the main incentive to improve the quality and completeness of data collections; and, more generally, educating target audiences on the public good nature of high-quality longitudinal education data. New systems often confront the challenge of being perceived mainly as auditing tools. It is thus important to increase awareness about the multiplicity of their potential uses beyond high-stakes evaluation. Strategies to prevent decontextualized uses of education data by the media can also help build support for new systems.

Session 5: Discussion and policy recommendations

46. In the closing session of the first day, participants divided into groups to discuss the presentations and reflect on the strengths and limitations of current systems, desirable functionalities and data for next-generation systems, and strategies to enhance stakeholder buy-in and use. Conclusions were then presented to other groups. The points below summarise the main themes in the discussion.

47. While many systems have made great progress from a technical point of view, most importantly by incorporating individual identifiers and expanding linkage possibilities across data elements and time points, a common impression across countries is that stakeholders often struggle to make a purposeful and efficient use of these sophisticated systems and the wealth of data they contain, a situation described with the metaphor “trying to drink from a fire hose”. This suggests that, for most systems, there is significant scope for improvement in customising functionalities and reporting for a greater range of user profiles. In this respect, education stakeholders have the responsibility to better define the purpose and key questions that education information systems should be designed to address, and to better identify the needs of their target audiences. Another reason for concern is that most current systems lack mechanisms to evaluate whether they are meeting their goals. These have to be formulated clearly so that the performance of any given information system can be assessed and improvement measures can be adopted when necessary.

48. A major impediment to using information systems to target specific school improvement initiatives is that data often lack contextual information, be it in the form of more and more diverse accompanying data elements or in the form of richer metadata about the data generation process. Teachers play a key role and systems could become more flexible in enabling practitioners to contribute data based on their professional judgement which can complement other assessment and evaluation measures. Systems should also become a support tool for teachers and include richer and timelier data to support instruction. For principals, information systems should also become early warning tools at the school level.

49. The need to rethink current models of professional development for data use was also discussed. Improving the usability of systems requires not only a greater degree of customisation to different users’ needs but also fostering their data literacy independently of (and prior to) opening up the use of data systems.

Day 2: Using next-generation information systems for continuous improvement in education

50. The second day of the workshop explored the different uses of data and the potential uses of next-generation systems to improve and innovate education: advancing and disseminating educational research; analysing and redesigning teaching and learning practices; and strengthening evaluation for school and system level improvement. Conditions and barriers for effective data use were also discussed.

Session 6: Supporting educational research

51. This session discussed the potential of longitudinal information systems to support educational research. Through the integration of data from multiple sources and points in time, longitudinal information systems open up new possibilities for educational researchers to analyse the determinants of student success as well as the effectiveness of educational innovations and reforms. Participants were also invited to discuss models of access that balance privacy concerns while enabling research that is useful for multiple audiences.

52. Chandra Haislet, Maryland Longitudinal Data System (MLDS) Center (United States), gave an overview of Maryland's longitudinal data system and discussed how it supports the state's research agenda. The system was launched in 2013 as a central repository for longitudinal data spanning from early childhood into the workforce that integrates data from three state agencies as well as various external sources. Its development owes to a leadership push from state policymakers seeking new tools to answer research and policy questions that required longitudinal data matches involving two or more agencies, mainly regarding student transitions out of secondary education. Public reporting on these questions is now carried out through featured dashboards on the system's website. More in-depth research is conducted by internal MLDS Center staff in partnership with the University of Maryland. Using tools directly built into the system, authorized researchers are able to view all database elements with the exception of individual IDs and other personal information. A specific initiative to strengthen research capacity is the MLDS Research Academy, which trains graduate students in state universities in using MLDS data within guided objectives. Based on this model, a specific training programme for policy decision makers will start in 2014.

53. Felice Levine, American Educational Research Association (AERA), argued that enhancing collaboration between the education research community and administrative data managers is essential for building cumulative knowledge about effective education. Data sharing facilitates the verification of hypotheses, the replication of studies and the testing of rival questions, to the benefit of a range of customers who, based on their own resources and expertise only, would not have the same capacity to advance knowledge on how to improve educational outcomes. The notion of collaborative customers may thus provide a rationale for data sharing partnerships and the definition of common data elements and standards. In line with a recent recommendation of the United States National Research Council, Levine also argued that, when dealing with administrative data containing personal private information (PPI), researchers should be required to adhere to confidentiality and data protection plans rather than to obtain the explicit consent of data subjects. When data collection involves interaction or intervention with subjects, though, informed consent agreement about the research and the data use should remain a requirement.

54. As data generation is rapidly transformed by technology and the interaction of public and private actors, there is an increasing need for practices that facilitate data use while protecting the identities of those who are part of administrative information systems and research data sets. Some of these practices include making available de-identified public use files for research and other purposes; the use of individual-level data under restricted condition (e.g., research data centres, third-party repositories, site licencing, supervised access); and statistical data alteration and simulation. A widespread adoption of these

practices could help expand access to valuable education data sources, which at present is often privatized to selected investigators. Education research could greatly benefit from expanded approaches to restricted access in keeping with privacy protection and data confidentiality rather than requiring explicit consent, at least when dealing with extant data.

55. Tom Bailey, Teachers College, Columbia University, provided a practical illustration of the importance of using longitudinal data to understand critical questions bearing on education policy. For instance, static, cross-sectional data have been shown to provide a vastly distorted view of the patterns of enrolment in community college in the United States, as only a tiny fraction of students actually follow canonical enrolment pathways¹. Before the possibility of tracking individuals longitudinally existed, this vast heterogeneity of student experiences was masked. Related projects have also challenged conventional wisdom about how university course credit accumulation influences individual labour market outcomes, and thereby led to changes in funding policies in higher education. Despite the proven benefits of constructing rich, longitudinal education databases, important barriers to data use for research purposes persist in the United States. Personal relationships between researchers and data managers can sometimes help overcome the reluctance to make data available, but the development of longitudinal information systems provides an opportunity to set up formal procedures to allow researchers to access, use and link these data which are independent of such relationships.

56. However, some positive developments with respect to data access occurred over the last decade. In the United States, federal grant programmes for the establishment of state-wide data systems played an important role in increasing possibilities for research, largely as they supported the emergence of institutional partnerships for data sharing. Strong leadership promoting collaboration between agencies was a common denominator of successful initiatives. Other important trends are the increasing use of register data and administrative records rather than new survey collections, and the growing acceptance of the culture of open data. The most promising models of data access to strengthen educational research appear to be those combining strict selection criteria of trusted users and little restrictions in terms of accessible data elements and linkage possibilities.

57. Drawbacks in access policies have also taken place, often motivated by commercial uses of school data or poorly contextualised media reporting. Stakeholders advocating a greater use of personal data for educational research have to date failed to argue their case as strongly as counterparts in other sectors, for example associations supporting medical research on rare diseases that requires extensive use of personal private information.

58. Another recurrent remark was that educational researchers should increase efforts to reach out to a broader set of audiences. While research is most often produced for policy makers and academics, many important educational decisions which research findings could positively inform are made by students, families and other stakeholders. Demands from different audiences may of course involve conflicting priorities, but broad research agendas and multi-layered dissemination strategies can help increase support for expanding access to education data for research purposes. A strategy to gain such support is to give greater opportunities to practitioners and families to provide inputs for the design of research agendas.

¹ Crosta, P. (2014), "Intensity and Attachment: How the Chaotic Enrollment Patterns of Community College Students Relate to Educational Outcomes", *Community College Review* (published online February 2014).

Session 7: Instructional practices and personalised learning

59. This session discussed the use of education data and longitudinal information systems to improve the personalisation of teaching and learning. The discussion focused on the possibilities offered by new data mining and learning analytics methods to identify individual learning patterns and tailor instruction. It also addressed the challenges of implementing these approaches in current school settings.

60. George Siemens, University of Texas, discussed models to improve personalisation based on insights from the emergent discipline of learning analytics. This research field applies a set of techniques such as relationship mining or knowledge domain modelling to a wide range of data reflecting learning processes, with the aim of understanding and optimising them. Specific applications range from trend analysis and prediction, to the provision of adaptive content and to structural analysis of learners' participation in social networks and information flows. The greater personalisation of learning would be achieved through individual learning plans that derive from mapping 'personal knowledge graphs' onto the concept architecture of a given field, with such representations of a person's actual knowledge being validated through the triangulation of evidence and feedback loops. Importantly, this can be applied not only to cognitive elements but also to learners' affective dimensions. A current challenge to this approach is to expand the scope of data capture on which these mappings are based, as it is to better understand how different types of data may reflect latent knowledge and skills. Nonetheless, learning analytics holds great promise to improve our ability to identify and assess learners' capacities by stitching together evidence from formal educational assessments and sequences of informal learning activities and interactions. There are thus great opportunities to link learning analytics applications and new longitudinal data systems.

61. Susan Fine, New Classrooms, described *Teach to One: Math* (TtO), a personalized learning model making an intensive use of data to tailor instruction to individual student needs. Implementation of the TtO model began in 2012 in eight schools in Chicago, New York City, and Washington, D.C. focusing on lower-secondary school mathematics. The objective of TtO is to provide instruction that is continually responsive to learners' abilities. The programme assesses students' skill levels on a daily basis and uses algorithms to target content delivery and assign students to varying instructional modes. These include teacher-led instruction, student collaborative work, virtual tutors and other educational software. The model relies on data from continuous formative assessment to identify individual learning gaps in skills maps, for instance on the mathematical relationships between ratios and rational numbers. Every day, students access computer dashboards displaying their progress and presenting them with tasks to work on their skills gaps as well as links to a variety of educational materials. Since skill maps are non-linear, students are allowed to move at their own pace and design their own "playlists" of tasks and skills. The large amount of data generated in this process is fed back to an underlying information system. The iteration of the model informs a daily reconfiguration of personalised learning paths and the design of broader, two-week instruction cycles. The process also provides teachers with real-time information about class and student performance through dynamic dashboards, allowing them to better target their support to students' learning.

62. The ensuing open discussion addressed a variety of issues. Many remarks invited a reflection on how developments in data mining and learning analytics could transform the role of teachers. It was argued that the affordances of technology in providing personalised feedback and disseminating content will lead to a redefinition of pedagogical and classroom management practices towards more immersive, active pedagogies and personalised tutoring. Moreover, if conventional class groupings were to blur as students gain some autonomy in setting their own learning paths, teachers may also need to develop a greater sense of shared accountability about student success.

63. The interoperability of existing longitudinal data systems with new software solutions for learning analytics and personalised instruction appears as an important need. The data repositories of

education agencies can furnish historical series for a wide array of variables and a wealth of contextual information to support benchmarking and statistical inference in these new applications. It is thus important that information systems at different levels are able to communicate and share data with each other, and that new learning analytics software allows connecting and enriching them with other data sources.

64. These innovative approaches to using data for educational improvement raise complex and still largely unresolved questions about privacy and data ownership. Privacy and ethics concerns feature prominently in current debates but changes in legal and governance systems have not followed the pace of developments in data use. Moreover, stronger evidence on the efficacy and cost-efficiency of learning analytics is needed before widespread implementation can be advocated.

Session 8: Innovative approaches to school and system improvement

65. This session discussed innovative approaches to the use of longitudinal information systems and education data more generally to improve school and system practices. Data collected within current systems can for instance support school self-assessment initiatives and strengthen public accountability by informing a better contextualisation of school performance.

66. In the first part of her presentation, Kim Schildkamp, University of Twente, summarised findings from a recent study on education data use in schools in five EU countries (The Netherlands, Germany, Lithuania, Poland and the United Kingdom)². This research shows that, despite the availability of a range of data sources, data use did only rarely result in school taking action to improve student learning. The United Kingdom (England) emerged as the only country in the study where information systems, tools and software to analyse data were commonly available in schools and where faculty collaborated actively around using data for instructional improvement. The school organisational factors having a strong influence on these practices were the presence of data experts on site, collaboration between teachers in the use of data, and training opportunities around data use.

67. The second part of the presentation elaborated on the importance of professional development for effective data use by describing the *Datateams* procedure developed in the Netherlands for using data for school improvement. Data teams consist of teachers, data experts and school leaders working together with a structured eight-step approach from defining a problem in the school, to formulating hypotheses, to making decisions based on data with regard to how to solve this problem, to evaluating the effectiveness of the measures adopted. At the core of the method is the objective of empowering schools to collect and analyse data to identify the causes of problems they confront. Ongoing research on the effects of the procedure in more than 40 Dutch schools indicates that this form of professional development is effective in increasing various dimensions of teachers' data literacy and skills for data use as well as teacher satisfaction³. The results also suggest positive effects on student achievement. The existence of school collaborative practices, access to high quality data and leadership support appear as critical conditions for the success of this type of initiatives.

68. Barbara Schneider, Michigan State University, discussed lessons from her work with the Michigan Consortium for Education Research (MCER), a partnership among the state of Michigan and two state universities to promote data-driven educational improvement at a systemic level. In line with its

² Schildkamp, K., Karbautzki, L. and Vanhoof, J. (2014). "Exploring Data Use Practices around Europe: Identifying Enablers and Barriers." *Studies in Educational Evaluation* 42: 15–24.

³ Ebbeler, J., Poortman, C.L., Schildkamp, K., and Handelzalts, A. (2014). "Effects of the data team procedure on data use". Paper presented at the 2014 AERA Annual Meeting, Philadelphia, PA.

mission to bridge research and practice in school settings, MCER is for example using historical records and gathering new data from Michigan high schools to better understand how course-taking patterns in high school influence student participation in higher education and to design interventions to boost the latter. The first step of the MCER approach is to meet with administrators and practitioners to identify key issues for improvement and build relationships and institutional trust. Work with data involves merging school data into larger analytical files that enable more detailed examination of the questions at stake. Essential to the process is the periodic provision of feedback through early descriptive reports and briefings so that educators can in turn provide input to fine-tune and align the research with their practical needs. The success of the partnership relies critically on the relational trust among all its stakeholders. This has been achieved by minimizing vulnerabilities, acknowledging contributions from all partners and sharing resources and training opportunities. Data sharing agreements clarifying data ownership and adopting strict rules to preserve privacy have also been of paramount importance. For instance, the privacy of individual records is protected through the use of computer-based Research Identification Codes (RICs) with no embedded meaning, while other personal identifiers are removed and stored separately from the files made available to researchers. MCER puts also a strong emphasis on its communication and dissemination activities with the legislature, the media and academic and general audiences.

69. The discussion then focused on the sustainability of data sharing and data use initiatives at the school and system level, especially with regard to funding and relational trust. Financial support from public agencies is common for most initiatives during their initial phase, but additional funding is often needed to secure continuity and should thus be planned for. However, the key to sustainability often rests on the success of internal capacity building, for instance by training data experts in schools so that data use practices can continue after interventions come to an end. Convincing schools and families of the relevance of this work is also important to gain trust and extended support; for this, interventions must address questions that are relevant for actors on the ground and not only academic debates.

70. The distinctiveness of the focus on data use vis-à-vis other initiatives for school improvement was also discussed. Participants agreed that data use must be integrated into existing practices rather than be taken as an aim in itself. Data remain a tool, and their value depends on how much they help improve decision-making.

Session 9: Conditions for effective data use

71. This session discussed the conditions that enable practitioners and decision-makers at the various levels to benefit from the wealth of information contained in longitudinal education data systems. The adoption of a data system is a socio-technological process of innovation and can involve (or be used to trigger) disruptions in established work and decision-making procedures.

72. Jeffrey C. Wayman, Wayman Services LLC., made a case for a social approach to the implementation of data systems. The experience of the United States over the last decade shows that many systems have failed to bring about data use that effectively supports instructional improvement. The dominant approach to rolling out data systems and promoting data use in schools focused narrowly on technology tools and largely neglected organisational aspects and user perceptions. The model was adopted on the premise that data use was self-evident after educators received training on how to utilise sophisticated tools for data retrieval and analysis, and that prior to the introduction of these systems teachers were not using information in any systematic way. In doing so, this approach failed to consider how data systems could disrupt teacher practices and presumed that new data would fill gaps in teacher knowledge. Most implementations following this model resulted in sporadic and poorly targeted use of the data systems in schools and in high levels of teacher mistrust.

73. A critical lesson from these experiences is that rolling out a data system involves an extended period of organisational adjustment, and that, once a certain level of technological sophistication has been achieved, success rests mainly on human components. Therefore, a social approach can be advocated which focuses first on building common understandings through direct communication and feedback from the targeted users, most importantly teachers, and on training for data use (understood as how to drawing meaning from data) rather than on system use⁴. Other important aspects are to embed data systems into collaborative practices in schools, and to promote a data culture which acknowledges that most often there are multiple valid answers to the complex questions that educators confront. Data systems must therefore become an empowerment tool which supports rather than replace teachers' professional judgment.

74. In the open discussion, participants debated evidence on different approaches to implementation. It was argued that, in most cases, high-stakes accountability has proven ineffective as a theory of action for data use, and that practitioners should be given the opportunity to express their views about desirable features of information systems without the framing of reporting requirements. Strategies to make information systems more attractive and relevant to educators therefore need to combine improvements in technical aspects and user-friendliness with professional development and protocols for effective data use. Participants also emphasized that data use should be promoted as a way of complementing (but not substitute other forms of) critical reflection about teachers' professional practice.

Session 10: Next steps toward solutions. Breakout sessions and final plenary discussion

75. In the closing session, participants reflected on the discussions held throughout the workshop to propose key messages and policy recommendations to promote the use of longitudinal data systems for innovation and improvement in education.

76. There was a call for greater international collaboration in order to establish a regular forum for the exchange of good practices in the use of information systems and education data at large. The OECD and other international organisations could take a lead role in putting forward a mission statement to gain attention from high-level bodies within countries and in coordinating working groups that help establish a common language to advance the conversation. Periodic meetings bringing together education policy makers, information system managers and data use experts are needed to engage as a community of practice and foster peer learning.

77. The OECD could also make an important contribution by providing an international perspective on the potential benefits and challenges associated with the establishment of education information systems. This analytical work should be framed within an innovation adoption model that acknowledges the varying pace of developments across countries and derives different messages for early adopters and for countries still in the process of building up their system infrastructure.

78. The accountability paradigm has been the key driver of developments in the establishment and use of information systems in education over the last decade. By and large, the period leaves a mixed balance of promising developments in data collection, data sharing and public reporting practices, on the one hand, and weak integration of data use with regular teaching and learning practices in schools, on the other. With the common objective of sustaining continuous improvement in education, the focus of the discussion is nonetheless gradually shifting from "evaluation" to "feedback for learning". Education

⁴ Cho, V. and Wayman, J. C. (2014). "Districts' efforts for data use and computer data systems: The role of sensemaking in system use and implementation". *Teachers College Record*, 116(2); Wayman, J. C. and Cho, V. (2014). "Realistic expectations in The Data-Informed District". Paper presented at the 2014 AERA Annual Meeting, Philadelphia PA.

agencies are therefore developing and extending their information systems with an increasing diversity of goals and approaches to data use.

79. With regard to a common vision for the future development of education information systems, other key messages were:

- Education policy makers should lead across-the-board stakeholder consultation to clarify the key questions that the development of information systems should contribute to solve. A clear formulation of purpose should precede the rollout of systems, which have often been put in the hands of practitioners without a common understanding of goals and functions.
- Information systems require formal and transparent governance structures. A special emphasis must be put on finding effective models for data ownership and data privacy protection, with the aim of facilitating data sharing and data use.
- Targeted professional development for educators should be an integral component of initiatives seeking to mainstream the use of information systems.
- In the era of “big data”, longitudinal information systems have the potential to become the basis for improved, more personalised learning practices. However, their development should not reinforce existing inequalities but rather help equalise opportunities for all.

Additional material

The workshop agenda and list of participants are included in the Annex to this document. They can also be found, alongside participant presentations, on the *Innovation Strategy for Education and Training* project pages on the OECD website⁵.

⁵ <http://www.oecd.org/edu/ceri/innovationstrategyforeducationandtrainingmeetingsandconferences.htm>

**ANNEX:
AGENDA AND LIST OF PARTICIPANTS**

Agenda

**“Fostering Innovation and Improvement in Education:
the Contribution of Longitudinal Information Systems”**

OECD Centre for Educational Research and Innovation (CERI)

30 June – 1 July 2014

New York City, United States

Barnard College, Columbia University, 3009 Broadway, New York, NY 10027

Background

As part of its effort to look at new ways to foster the innovation process in education, the CERI project “Innovation Strategy for Education and Training” is undertaking an international mapping and analysis of longitudinal data systems in education. This work explores the benefits that these knowledge management tools can bring to a variety of stakeholders in education. These include a greater personalisation of instruction to enhance learning; better and timelier feedback on performance and interventions; comparison and analysis tools to identify and tackle educational issues (e.g. dropout, gaps in achievement) through communities of practice; advancements in and better dissemination of educational research; easier evaluation and accountability practices; and efficiency gains in administrative operations.

The meeting will build on the dialogue started at the 2010 OECD-SSRC Workshop and take stock of developments in the field over the last four years. It will also serve to present results from the CERI survey of information systems in education which covers over 60 systems in more than 25 OECD and partner countries.

Objectives

The aim of this international meeting is to foster discussion about the latest developments in the design of longitudinal education data systems, the conditions under which these tools are most effective, and the main challenges associated with their use. It will convene about 40 policy makers, data use experts, system developers and practitioners in the United States, other OECD countries and partner economies to discuss the potential of “big data” for educational improvement and innovation and the contribution of next-generation education data systems.

The working language of the workshop will be English.

DAY 1 – MONDAY 30 th JUNE	
MAPPING TECHNICAL FEATURES, CURRENT USES AND OPPORTUNITIES FOR DEVELOPMENT	
8.30–8.50	Registration and coffee
8.50–9.00	Welcome
9.00–12.00	Session 1: The state of longitudinal information systems in education today
<p>This session will outline the objectives, characteristics and functionalities of existing education information systems across the OECD area and beyond. It will start with an overview of the results of the CERI survey, followed by short presentations of specific systems in various OECD countries. The objective will be to establish where most information systems currently stand, and to highlight interesting features and uses as well as directions for future development.</p>	
9.00-9.30	<p>An international mapping of longitudinal information systems in education. The CERI survey <i>Speakers: Stéphan Vincent-Lancrin and Carlos González-Sancho, OECD</i></p>
9.30-10.00	<p>Presentations of information systems in the US and other OECD countries <i>Moderator: Nancy Sharkey, National Center for Education Statistics</i> <i>Speakers: Tadesse Haile, Education Statistics and Analysis Branch, Ontario</i> <i>Kari Steenstrup, Directorate for Education and Training, Norway</i></p>
10.00-10.30	<i>Coffee break</i>
10.30-12.00	<p>Presentations of information systems (continued) and open discussion <i>Speakers: Melissa Beard, Washington State Education Research and Data Center</i> <i>Frank Barnes, Charlotte-Mecklenburg Schools</i> <i>Birgit Lao-Peetersoo, Ministry of Education and Research, Estonia</i></p>
12.00-13.00	<i>Lunch break</i>
13.00-14.00	Session 2: Information systems in higher education and links with the labour market
<p>This session will focus on the characteristics and potential of longitudinal information systems in higher education, which are typically developed independently from systems for K-12 schooling. Participants will discuss challenges and opportunities given the greater degree of autonomy of higher education institutions, the greater mobility of students across programs and institutions, and the increasing importance of life-long learning. Linkages between higher education and labour market information systems to better match the supply and demand of skills will also be addressed.</p> <ul style="list-style-type: none"> • Links between K-12 and higher education information systems • Links between higher education information systems and labour market data • Providing information to students and families for programme choice in higher education • The role of employers: What access policy? What data can they feed back? • Using data to identify professional development needs and support life-long learning <p><i>Moderator: Tom Bailey, Teachers College</i> <i>Speakers: Nara Vieira, National Institute for Educational Research, Brazil</i> <i>Daniel Domagala, Colorado Department of Education</i> <i>Regina Radinger, Statistics Austria</i></p>	

14.00-15.00	Session 3: Roundtable next-generation systems. Promising features in the R&D pipeline
<p>This session will invite designers, vendors and entrepreneurs to present promising ideas about what next-generation systems could look like in the coming years by incorporating technological innovations in the R&D pipeline as well as new business models to overcome the traditional institutional barriers around data use.</p> <p><i>Moderator: Hui Soo Chae, EdLab, Teachers College</i></p> <p><i>Speakers: Gerri Burton, New Learning Ventures</i> <i>Nelson Gonzalez, Declara</i> <i>Chris Lohse, Renaissance Learning</i></p>	
15.00-15.30	<i>Coffee break</i>
15.30-16.30	Session 4: Setting up a new system: policy mandates and practical considerations
<p>This session will give education agencies in the process of setting up or redesign their information systems an opportunity to discuss their policy mandates as well as the practical and technical challenges they confront. The discussion will invite reflections on planning strategies from different political and institutional contexts.</p> <p><i>Moderator: Tate Gould, AnLar Inc.</i></p> <p><i>Speakers: Teresa Evaristo, Ministry of Education, Portugal</i> <i>Agustín Caso-Raphael and Arturo Cervantes, National Institute for Educational Evaluation, Mexico</i></p>	
16.30-17.15	Session 5: Discussion and policy recommendations (breakout)
<p>In this session, participants will reflect in small groups on the technical strengths and limitations of today's longitudinal information systems and the horizon for the adoption of new features: is there a shared set of design specifications across countries? What room do they leave for extending the core model? What are the key factors driving replacement and then adoption of new developments? Is there an appetite for new features among the key stakeholders? Participants will formulate policy recommendations for systems to communicate across countries and jurisdictions, and to help those who are in the process of establishing new systems.</p>	
17.15-18.00	Session 5: Discussion and policy recommendations (plenary)
<p>In this session, we will continue the discussion in plenary and debate the recommendations of the different groups.</p>	
18.00	<i>Cocktail</i>

DAY 2 – TUESDAY 1st JULY

USING NEXT-GENERATION INFORMATION SYSTEMS FOR CONTINUOUS IMPROVEMENT IN EDUCATION

8.30–8.45

Coffee

8.45-10.00

Session 6: Supporting educational research

This session will discuss the potential of longitudinal information systems to support educational research and help disseminate research findings to the benefit of a wider range of stakeholders. Through the integration of data from multiple sources and points in time, longitudinal information systems open up new possibilities for educational researchers to analyse the determinants of student success as well as the effectiveness of educational innovations and reforms. In return, next-generation information systems can benefit from educational research by incorporating new evidence and insights generated by research activity. For researchers, access to longitudinal data systems can sometimes be difficult in practice and we will discuss models that manage to balance privacy concerns and flexible access to enable research that is useful for a broad audience.

- Data-sharing agreements, access protocols and conditions to link multiple data sets
- Integrating research findings and recommendations into information systems
- Knowledge sharing practices: conversations outside the research community

Moderator: Barbara Schneider, Michigan State University

Speakers: Chandra Haislet, Maryland Longitudinal Data System Center

Felice Levine, American Educational Research Association

Tom Bailey, Teachers College

10.00-11.15

Session 7: Instructional practices and personalised learning

This session will discuss the use of education data and longitudinal information systems to better tailor instruction to student learning needs. The discussion will focus on practices at the classroom and school levels and explore the content and forms of feedback that educators need to identify problems and adjust instruction in effective ways.

- Data generation through the use of classroom devices and formative assessment tools
- Improving the speed and quality of feedback to teachers
- Data mining and learning analytics: identifying learning patterns to inform better practice
- Integrating learning management platforms with longitudinal information systems
- Integrating early warning and intervention systems with longitudinal information systems
- The use of information systems for personalised paths in higher education

Moderator: Jennifer Davis Poon, The Council of Chief State School Officers

Speakers: George Siemens, University of Texas

Sue Fine, New Classrooms Innovation Partners

11.15-11.30

Coffee break

11.30-12.45

Session 8: Innovative approaches to school and system improvement

Data collected within current systems are often used to increase the efficiency of administrative and managerial operations, but also for evaluation and accountability purposes. This session will discuss the different roles that longitudinal information systems play and could play to monitor educational goals and policies but also help build consensus and capacity at multiple levels.

- Strengthening public accountability through a better contextualisation of school performance
- School self-assessment
- The high-stake use of assessment and information systems (teacher value-added, school ratings, etc.)
- Innovative models to enhance trust/acceptability within systems

Moderator: Jennifer Davis Poon, The Council of Chief State School Officers

Speakers: Kim Schildkamp, University of Twente, The Netherlands

Barbara Schneider, Michigan State University

12.45-13.45	<i>Lunch break</i>
-------------	--------------------

13.45-15.00	Session 9: Conditions for effective data use: what is the evidence?
<p>This session will present and discuss some studies about the use of data within schools or districts to cast light on the conditions that enable practitioners and decision-makers at the classroom, school, district and state/region level to benefit from the wealth of information contained in longitudinal education data systems. The adoption of a data system is a socio-technological process of innovation and can involve (or be used to trigger) disruptions in educators' work routines. Presentations will address the role of longitudinal information leveraging both existing expertise and enabling new conversations. It will also discuss the modes of work organisation that facilitate the use of data systems for school improvement. Challenges to effective use will also be addressed, including educators' time constraints, equipment inadequacies in schools, limited data literacy skills, and mistrust about data use derived from a high-stakes data culture that has traditionally focused on compliance rather than improvement.</p> <ul style="list-style-type: none"> • Developing data literacy among educators: what professional development to overcome skills gaps? • Developing organisational routines that facilitate the use of data systems for improvement • Effective visualization and reporting tools: ease of use, relevance and customisation possibilities • Enabling new conversations and ways of collaboration through discussion protocols and platforms • What access-to-data policies are needed to support fruitful conversations? • Developing a positive data culture <p><i>Moderator: Felice Levine, American Educational Research Association</i> <i>Speaker: Jeffrey Wayman, Wayman Services</i></p>	

15.00-15.15	<i>Coffee break</i>
-------------	---------------------

15.15-16.30	Session 10: Next steps toward solutions. Breakout sessions and final plenary discussion
<p>In the closing session, participants will first discuss in groups to propose key messages or policy recommendations to facilitate the use of longitudinal data systems for innovation and improvement. Discussants will put forward their views on what the focus should be for policy actions and entrepreneurial efforts over the coming years. Proposals will be then shared and discussed in a final plenary session that will outline recommendations for next-generation data systems to effectively impact on education research, evaluation and planning, and teaching and learning practices.</p> <ul style="list-style-type: none"> • What design features should be improved/adopted in existing/new information systems? • What are the next steps to effectively diffuse good practices in the use of information systems and education data at large? • What avenues exist for international action and collaboration? 	

16.30	<i>Close</i>
-------	--------------

17.00 – Visit to EdLab (optional)
EdLab is a research, design, and development unit at Teachers College, Columbia University. EdLab has developed numerous projects around education data collection and analysis making use of new technology tools.

List of participants

**“Fostering Innovation and Improvement in Education:
the Contribution of Longitudinal Information Systems”**

**30 June – 1 July 2014
New York City, United States**

Austria	Ms. Regina RADINGER Statistics Austria
Brazil	Ms. Nara Núbia VIEIRA National Institute for Educational Research (INEP)
Canada (Ontario)	Mr. Taddesse HAILE Education Statistics and Analysis Branch
Estonia	Ms. Birgit LAO-PEETERSOO Ministry of Education and Research
Mexico	Mr. Agustín CASO-RAPHAEL National Institute for Educational Evaluation (INEE) Mr. Arturo CERVANTES National Institute for Educational Evaluation (INEE)
Netherlands	Ms. Kim SCHILDKAMP University of Twente Ms. Pauline THOOLEN Ministry of Education, Culture and Science
Norway	Ms. Kari STEENSTRUP Directorate for Education and Training
Portugal	Ms. Teresa EVARISTO Directorate General for Education and Science Statistics
Slovenia	Mr. Gašper CANKAR National Examinations Centre
Spain	Mr. Valentín DUEÑAS JIMÉNEZ Ministry of Education
United States	Mr. Tom BAILEY Teachers College, Columbia University Mr. Frank D. BARNES Charlotte-Mecklenburg Schools, North Carolina

Ms. Melissa BEARD
Washington State Education Research and Data Center

Ms. Gerri BURTON
New Learning Ventures

Mr. Hui Soo CHAE
Teachers College, Columbia University

Ms. Jennifer DAVIS POON
The Council of Chief State School Officers

Mr. Daniel DOMAGALA
Colorado Department of Education

Ms. Susan FINE
New Classrooms Innovation Partners

Mr. Nelson GONZALEZ
Declará

Mr. Tate GOULD
AnLar Inc.

Ms. Chandra HAISLET
Maryland Longitudinal Data System Center

Mr. Robert HUGHES
New Visions for Public Schools

Ms. Kathryn KOLBERT
Barnard College, Columbia University

Ms. Felice LEVINE
American Educational Research Association

Mr. Christopher LOHSE
Renaissance Learning

Ms. Barbara SCHNEIDER
Michigan State University

Nancy S. SHARKEY
National Center for Education Statistics

Mr. George SIEMENS
University of Texas Arlington

Mr. Jeffrey C. WAYMAN
Wayman Services, LLC

Mr. Daniel YAO
TraceRoute

OECD

Mr. Carlos GONZÁLEZ-SANCHO
Directorate for Education and Skills

Mr. Stéphan VINCENT-LANCRIN
Directorate for Education and Skills