e-assessment community

7TH FLIP+ ANNUAL EVENT, LISBON, 19TH – 21ST JUNE 2024

Venue: Secondary School Pedro Nunes, Av. Álvares Cabral 39, 1250-015, Campo de Ourique, Lisbon

AGENDA

Day 1 will start with welcomes, then we will hear from member institutions and associated contributors who will share their e-assessment experiences. In the afternoon there is the first of the four workshops followed by discussion sessions on the FLIP+ working Groups and the International Item Library.

Day 2 will include three interactive workshops, followed by a closing plenary session to wrap the activities of the whole event, which will end at 16:00.

See pages 3–9 for details on the scheduled sessions and pages 10–12 for details of presenters.

Wed 19th June

10:00 – 16:00	FLIP+ Working Groups meeting (only for FLIP+ Working Group members), IAVE Limoeiro, Travessa das Terras de Sant'Ana, 15, 1250-269, Campo de Ourique, Lisbon
18:00 – 18:30	On-site registration and collection badges
18:30 – 20:00	Welcome reception supported by IAVE Limoeiro, at Secondary School Pedro Nunes, Av. Álvares Cabral 39, 1250-015, Campo de Ourique, Lisbon

DAY 1: Thu 20th June

08:30 onwards	On-site registration and collection badges
09:15 – 10:00	Welcome address · To IAVE – *Luis Santos, Director, IAVE* · To FLIP+ Annual event – *Thierry Rocher, President of FLIP+* · To the event programme – *Amina Afif, Executive Secretary, FLIP+*
10:00 – 10:45	Member presentations POR: *National Institute for Educational Evaluation (IAVE): The Digital transformation on large scale assessment in Portugal: What did we learn? Luis Santos and Rui Pires* UK: AlphaPlus - *Transitioning to on-screen assessments in Wales, Ben Rockliffe*
10:45 – 11:15	Group Photo and Coffee break
11:15 – 12:30	Member presentations (*contd.)* BRA: *CAEd - Challenges and solutions in assessing students in Brazil, Carlos Palacios* FRA: *DEPP – Student assessment in France: recent developments, challenges and opportunities, Anaïs Bret* SWITZ: *SRED - E-assessment in Switzerland: where do we stand? A brief overview, Franck Petrucci* LIT: *NAE – Lithuanian experience in assessing competences in e-testing format in high-stakes examinations, Asta Ranonyte*
12:30 – 13:30	Lunch break
13:30 – 15:00	Unlocking the Potential of AI in assessment: A Deep Dive with CitoLab NED: *CitoLab Presenters - Dr. Joost Kruis, Eva de Schipper and Tjeerd Hans Terpstra*
15:00 – 15:30	Coffee break

15:30 – 17:00 FLIP + Working Groups and International Item Library Update

Information about the progress in setting up the IIL project and open discussion with the FLIP+ community members to align their needs with the goals of the IIL

Facilitators: *Thierry Rocher and Amina Afif*

17:00 End of DAY 1 sessions

19:00 – 21:00 Reception Cocktail: Miradouro de Baixo - Carpintarias SL Rooftop

- *https://www.carpintariasdesaolazaro.pt/miradouro-de-baixo*

DAY 2: Fri 21st June

09:00 – 10:30 AI and the future of assessments

- *CitoLAB (NED) Presenter: Joost Kruis / Tjeerd Hans Terpstra*
  Can you use AI to grade final exams?
- *Vretta (CAN) Presenter: Vali Huseyn*
  AI-Enhanced Strategies for Collusion Detection and Prevention in Assessment Administration
- *CUPA – DEFI (UK) Presenter: Sarah Hughes, Bryan Maddox*
  The Futures of Assessment: navigating uncertainties through the lenses of anticipatory thinking

10:30 – 11:00 Coffee break

10:30 – 12:30 Workshop: AI and assessment design

- *University of Stavanger - Norwegian Centre for Reading Education and Research (NO) Presenter: Bente R. Walgermo PhD*
  Developing a Human Reading Assessment – an AI-assessment project
- *University of Stavanger - Norwegian Centre for Reading Education and Research (NO) Presenter: Per Henning Uppstad*
  School entry detection of reading difficulties using machine learning on gameplay process data
- *AQA (UK) Presenter: Lucy Howarth and David West*
  Exploring how technology could mitigate assessment material errors

12:30 – 13:30 Lunch break

13:30 – 15:00 Workshop: AI and marking

- *Cambridge University Press & Assessment (UK) Presenter: Sanjay Mistry and Dr. Jesse Dvochak*
  Using AI Assisted auto-marking with a range of item types- a Proof of Concept
- *DIPF (DE) Presenter: Ulf Kroehne et al., DIPF | Leibniz Institute for Research and Information in Education*
  Digital Formative Assessment for Learning ("Alea.schule")
- *CAEd (BRA) Presenter: Jairo Francisco de Souza, Roberta Melo*
  Use of AI to correct audios for the assessment of reading fluency

15:00 – 16:00 Closing session and end of event

FLIP+ 7TH ANNUAL EVENT 2023: DAY 1 SESSIONS

MEMBER PRESENTATIONS

IAVE (POR)

Title: The Digital transformation on large scale assessment in Portugal: What did we learn?

Presenters: Luis Santos and Rui Pires

IAVE has been undertaking a digital transition process in large scale national assessment in recent years. This process, considering its complexity, has brought great challenges to the educational system. From network infrastructures and the diversity of computer equipment in schools, as well as keyboard writing in constructed items and symbolic algebraic writing in mathematics and science, many risk factors occur when implementing a project of this nature.

In this presentation, we will address this digital transition in the Portuguese educational system, as well as its advantages and challenges.

AlphaPlus (UK)

Title: Transitioning to on-screen assessments in Wales

Presenter: Ben Rockliffe

In Wales, all children aged 7 to 14 take national assessments in procedural numeracy (number, measuring and data skills), reading and numerical reasoning (solving problems). In the past these assessments were taken entirely on paper, but Welsh Government decided to move them completely on screen in a phased roll out starting in 2018.

AlphaPlus leads this multi-partner project team on behalf of Welsh Government, which is a major statutory rollout in UK schools. It provides one vision for how the requirements of national assessment can be aligned with the real teaching and learning needs of teachers and pupils.

Our presentation will talk about our experience of building and sustaining a user-led assessment programme, and will cover the following key features of our innovative approach: working with and preparing schools for implementation of onscreen assessments; constructing rich and insightful reports to help teachers, learners and parents understand what they need to do next to improve competence; designing an interface that will deliver an engaging experience – including for children as young as seven, who will use diverse devices and operating systems; running the system over multiple years – producing annual reports for government and public consumption; in particular, using our exceptional data set to make important statements about the impact of Covid on learning in Wales, specifically patterns in reading and numeracy attainment: from 2018/19 to 2022/23; and working with teachers and learners on an ongoing basis, to understand their perceptions and amend assessment designs in response to these.

CAED (BRA)

Title: Challenges and solutions in assessing students in Brazil

Presenter: Carlos Palacios

In 2024, CAEd/UFJF assessed students from more than 20 Brazilian states, facing the challenge of producing data that can be analysed by managers and teachers to improve learning in different regions and contexts. In this sense, we will present some of the main challenges faced and solutions found by CAEd and Education Departments for applying tests, producing data and developing feedback that support policies and practices to guarantee the right to learn for students across Brazil.

DEPP (FRA)

Title: Student assessment in France: recent developments, challenges and opportunities

Presenter: Anaïs Bret

The presentation aims to highlight recent advancements in standardized student assessments in France, particularly focusing on digital assessments. Significant progress has been made in these assessment programs. By September 2024, all students from grades 1 to 10 — approximately 8 million students — will be assessed in French and Mathematics, with half of them taking the tests on computer. These assessments serve both individual diagnostic and school monitoring purposes. The presentation will explore the challenges and opportunities of digital assessments, including test mode comparability, reporting systems, accessibility issues, and potential innovative approaches.

SRED (CH)

Title: E-assessment in Switzerland: where do we stand? A brief overview

Presenter: F. Petrucci, C. Denecker, O. Prosperi & E. Roos (Service de la recherche en éducation – Canton of Geneva)

In Switzerland, the organization of compulsory education is the responsibility of the cantons. Following the approval of the constitutional articles on education by the Swiss people in 2006, the cantons are required to harmonize some key features of their school systems and ensure the quality of the Swiss education system. The work carried out by the Swiss Conference of Cantonal Ministers of Education (EDK) led to the Intercantonal Agreement on Harmonisation of Compulsory Education in 2007 (HarmoS Agreement) and the adoption of national educational standards in 2011. These standards, also known as « core competencies » (UGK) correspond to the minimum competencies that all Swiss students are expected to master by the end of each school cycle (4th, 8th and 11th grades) in four domains: language of instruction, mathematics, foreign languages, and science. This communication shows how large-scale e-assessments are used to verify that these core competencies have been achieved.

National Agency for Education (NAE) (LIT)

Title: Lithuanian experience in assessing competences in e-testing format in high-stakes examinations

Presenter: Asta Ranonyte

In one year, Lithuania has experienced two major challenges in its education system: the launch of the Curriculum for the Development and Assessment of Student Competences and the simultaneous organisation of high-stakes matriculation exams in an e-testing platform. Systematically preparing for the changes in external assessment, we developed tasks in an electronic environment and invested a lot in the richness of the tasks, which was illustrated by the different formats (texts and media, different levels of interactivity); the contexts, which required selecting information, evaluating it, grouping it; and the actions required to perform them, which involved thinking procedures of different complexity. After the e-testing campaign we received feedback from the teacher community that MC and complex items with multiple constructed responses do not measure competences, but only test basic knowledge. The main component of the students' responses consisted of reflections on the emotional strain experienced. I would like to share some examples and insights on what lessons we have learnt from the intensive high stakes exams of 20 subjects e-testing campaign in 2024, and what makes us continue with the planned activities despite the challenging consequences for the education system in Lithuania.

WORKSHOP: Unlocking the potential of AI in assessment: a deep dive with CitoLab

CitoLab (NED)

Title: Unlocking the Potential of AI in assessment: A Deep Dive with CitoLab

Presenters: Dr. Joost Kruis, Eva de Schipper and Tjeerd Hans Terpstra

In recent years, the explosion of generative Artificial Intelligence (genAI) has had a profound impact on educational assessment and learning, unleashing a new era of possibilities and innovations. From personalized learning paths to advanced grading systems, AI is revolutionizing the way that we understand and implement assessments. Cito's research and innovation department CitoLab is embracing this development and using its own expertise to identify where AI can add value.

During our session, we will delve into the specifics of how Cito is harnessing AI (genAI as well as 'traditional' AI) within key phases of the testing cycle: construction, test administration, scoring and reporting. We will illustrate our journey with a demonstration of some of our prototypes and share the insights we have gained on these topics:

- How can item construction be assisted by AI?
- How can we achieve higher learning gains by using AI to recommend practice items?
- How can AI contribute to quick and consistent scoring?
- How can we use AI to enhance the value of rubrics in formative assessment?

Join us as we take a deep dive into the exciting world of AI in education and assessment. Using AI in a careful and deliberate way, we want to explore how we can pave the way for a smarter and efficient educational framework.

DISCUSSION: The FLIP+ Working Groups and the International Item Library (ILL)

FLIP+ and IIL

Title: The Role of the FLIP+ working Groups in the development of the IIL

Presenters: Amina Afif and Thierry Rocher and members of the FLIP+ Working Groups

In this session, the objective is to share information about the preparation of the launch of the International Item Library (IIL) project and to align the needs of the FLIP+ community with the goals of the IIL.

Participants will discover their how their institution could potentially contribute to the IIL's creation, benefit from its offerings, and shape its development. Be part of this interactive discussion and help shape the IIL's journey towards empowering global education collaboration.

AI and the Future of Assessment

CITO LAB (NED)

Title: Can you use AI to grade final exams

Presenter: Dr Joost Kruis / Tjeerd Hans Terpstra

In this session, we seek to explore the intriguing concept of AI systems grading final exams autonomously—a parallel drawn from fully self-driving cars. Drawing on this analogy, we delve into the ethical implications surrounding high-stakes and low-stakes assessment scenarios. Inherent within these discourses are concerns about the opacity of AI – it's not human, and its potential bias in its decision-making processes (as well as lack thereof!) is currently less predictable. As AI's grading proficiency surpasses human graders in controlled conditions, a critical question arises: at what point of accuracy do we start placing our trust in AI over human graders? This session intends to ignite discussions and debates around these key issues and to contribute to an ongoing conversation about the future of AI in educational assessments.

Vretta (CAN)

Title: AI-Enhanced Strategies for Collusion Detection and Prevention in Assessment Administration

Presenter: Vali Huseyn

The rapid integration of Artificial Intelligence (AI) in educational assessments introduces unprecedented challenges and opportunities, particularly in maintaining fairness and objectivity in administration of assessments. This presentation's focus is on the dual roles of AI in enhancing and potentially compromising assessment integrity, focusing on collusion detection and prevention in assessment administration. Highlighting the slower institutional adoption compared to the swift candidate utilization of generative AI tools, it underscores the essential balance required in navigating the new reality of human-AI collaboration. Practical recommendations for stakeholders, including policymakers, educators, and guardians, are outlined to address specific scenarios of AI use and misuse. This presentation aims to explore the transformative impact of AI in educational assessments, advocating for a future where AI's capabilities are leveraged responsibly to uphold the highest standards of integrity and fairness.

CUPA – DEFI (UK)

Title: The Futures of Assessment: navigating uncertainties through the lenses of anticipatory thinking

Presenter: Sarah Hughes and Bryan Maddox

We propose to present the findings of collaborative research project conducted in 2023 by Cambridge University Press and Assessment, the Digital Education Futures Initiative, and University of Cambridge partners. The aim of the project (trailed at least year's Flip-plus event) was to anticipate assessment in 2050, and to use Futures Thinking methods to identify and appraise plausible signals for the future change. The completed research report (which we will make available), identified four major drivers for change: climate change; AI tutors and personalised learning; immersive, interactive and augmented assessments; and the promotion of human flourishing and wellbeing. We then "back cast" from 2050 to the present, to describe five stages of activity starting now with a phase of research and awareness followed by pilot and development, adoption and transition, expansion and collaboration leading us to full integration and enhancement. This work reminds us that future is not inevitable or fixed but is the product of human agency: a combination of thought, decisions, actions and reactions and provides some insights into what those actions could be to help us work towards our preferred assessment future. We will discuss our next steps, and give the Flip-plus participants the opportunity to ask questions, and to explore what these drivers might mean for their own assessment contexts.

AI and Assessment Design

University of Stavanger - Norwegian Centre for Reading Education and Research (NO)

Title: School entry detection of reading difficulties using machine learning on gameplay process data

Presenter: Njål Foldnes, Per Henning Uppstad, Steffen Grønneberg, and Jenny M. Thomson

Reading Difficulties (RD) are among the world's best-documented mental disorders. Early detection of RD may enhance life quality of these vulnerable citizens. However, despite huge advances made, the best current screening paradigm is failing to identify over 50% of the school starters who will fall behind in reading. In this study we take an extra-paradigmatic approach to screening for RD utilizing machine learning analysis on school starters' progress data from playing a literacy game 10 minutes a day for five weeks. Models were trained on the process data combined with results from the end-of-year national screening test. The best machine learning models correctly identified 75% of the students at risk for developing reading difficulties. Consequently, the present study documents an approach for early detection of RD that by far surpasses the achievements of the current paradigm.

University of Stavanger - Norwegian Centre for Reading Education and Research (NO)

Title: Developing a Human Reading Assessment – an AI-assessment project

Presenter: Bente R. Walgermo PhD

In the context of AI and digital transitions in education, we will present our ongoing AI-assessment project called HUMAN reading assessment, which explores how we can develop a HUMAN digital reading assessment that presents students with tasks adapted to their interests and level of reading skill.

Artificial intelligence (AI) is entering schools at full speed, bringing with it both opportunities and challenges. Human Reading Assessment (HUMAN) aims to explore how technology can be applied in a way that places the student in the centre. During their years in school, students must undertake numerous reading assessments. These assessments form the basis of HUMAN, which seeks to respond to the challenges of AI on several levels: At the most basic level, HUMAN will develop and investigate algorithms, and explore how these can be applied in human centered reading assessments. The next level involves examining the content of the learning technology. In this case, the content are the texts and assignments found in the assessment. HUMAN will investigate how learning technology can contribute to designing texts and tasks that are both engaging and supportive of learning. In addition, the project aims to investigate how assessment results should be presented, in order for teachers to use them to adapt their instruction. At its most general level, HUMAN investigates what the application of AI in learning technology means for children's experience of the world and their agency in it, as well as its consequences for children's rights. How AI is implemented and presented in schools has major consequences for the individual's ability to act independently later in life. HUMAN is therefore also about examining the role of the education system for the individual, and for society and working life. The knowledge developed in the project will provide important input for future policy-making. Findings from HUMAN will also be directly applicable for school administration on all levels, for teachers and other actors in the education sector. HUMAN will take place in close collaboration with Norwegian schools. It is an interdisciplinary project, involving researchers in literary studies, literacy and reading studies, computer science and law.

AQA (UK)

Title: Exploring how technology could mitigate assessment material errors

Presenters: Lucy Howarth and David West

The increased demand on subject matter experts during exam seasons provides a strong case for exploring how new technology could quality assure and check assessments, to reduce dependency on these individuals. Possibilities include machine translation for checking text in modern foreign language assessments, copyright checking, and the use of generative AI for checking questions relating to methods and code in computer science.

While these tools may have great potential, there are some important limitations of AI. Examples include that its capability to generate natural language across many academic subjects carries a risk of falsehoods through hallucination. It is not able to identify any instance where the description of an external source is at fault, or where the wrong source has been assigned for the task. It would need support with particularly specialised vocabulary. It would struggle to compare complicated mathematical equations for their mutual consistency. Understanding the scope of a mathematical task or interpreting diagrams seem beyond reach for now.

The quality of assessment tasks remains the responsibility of awarding organisations, and this cannot be delegated to AI. Technology can support the quality control of assessments, but it cannot replace the role of human experts in high-stakes assessment production.

AI and Marking

Cambridge University Press & Assessment (UK)

Title of presentation: Using AI Assisted auto-marking with a range of item types- a Proof of Concept

Presenter: Sanjay Mistry and Jesse Dvochak

Within Cambridge University Press and Assessment, there is the desire to integrate AI based technologies to support both internally to the organisation but also where it can add value to the educational landscape, particularly in relation to assessment. There are many projects looking at the use of Generative AI and other AI applications. As part of this, a team from our Digital High Stakes programme are exploring, the accuracy of marking a range of item types using a third-party AI powered automaker. We would like to share with delegates the work we are doing in this space through the proof of concept. During the presentation we will discuss the type of items selected across a range of assessments, the approach to training the algorithm and number of scripts required, as well as security management. Following this, we will talk through the outcomes and how we measured the accuracy of the AI marked responses

DIPF (DE)

Title: Digital Formative Assessment for Learning ("Alea.schule")

Presenter: Ulf Kroehne et al., DIPF | Leibniz Institute for Research and Information in Education

For many years, the open-source research software "CBA ItemBuilder" was mainly used in large-scale assessments. However, as we showcase in a recent development, the technology can also be used for interactive, digitally-based formative assessments for everyday classroom use. Using items available as Open Educational Resources (OER), we computerized 45 short tests in mathematics, with immediately available feedback for teachers. In this contribution, we share our insights and experiences of the journey from the first idea to the launch of the research-practice transfer product. The presentation will cover how we developed items that contain appealing interactive components suitable tablet-based math assessment, our motivation and first results of experiments to integrate speech recognition and speech-to-text conversion into the items, the way we aim at transparency for teachers using a replay complete

collection of log data, NLP/transformer-based assistance for the coding of open-ended text responses and AI applications for coding graphical responses, the generation of meaningful feedback and the presentation of feedback in a learning analytics dashboard for teachers, implemented into our platform.

CAEd (BRA)

Title: Use of AI to correct audios for the assessment of reading fluency

Presenter: Jairo Francisco de Souza and Wagner Silveira Rezende

Since 2018, CAEd/UFJF has been conducting reading fluency assessments with education systems and departments, to contribute to literacy at the right age. The reading fluency assessment consists of recording students' reading to be analysed by teachers and identified according to reader profiles. In 2023, artificial intelligence technology was implemented, to reduce correction costs and make the process faster. In this sense, it will be presented how CAEd/UFJF built this technology and what are the achievements, the obstacles, and the future perspectives.

PRESENTERS, PANELISTS AND CONTRIBUTORS

Amina Afif is an experienced project coordinator of national and international multilingual teams with a demonstrated history of working with policymakers, researchers and practitioners in the field of educational policy, evaluation, digital assessment, data use and school improvement. Worked with the Ministry of Education in Luxembourg for 15 years, heading the division for the development of school quality and data use for school improvement. Currently working as an independent consultant as an education policy advisor, and transformational life coach. She is a Founding member and Executive Secretary of the FLIP+ e-assessment association.

Anaïs Bret is the Deputy Head of the Office for Student Assessments at the DEPP and a member of the Steering Committee of the International Item Library consortium. As a former teacher, she has taught physics to middle school students for 10 years. She was also in charge of standardized science assessments in France, in particular PISA science and TIMSS science. She became the coordinator of large-scale assessment programs in secondary school levels in 2019.

Asta Ranonyte is the Deputy Director of the National Agency for Education in Lithuania. She is responsible for horizontal and mutual interconnections between the content, its adaption to students with special needs and assessment of the educational process as well as data use to improve student achievements. Her previous position was the Head of the Division for Developing Achievement Assessment Tools. She worked in the field of assessment for more than 10 years. During the last years, she took part in the processes related to implementation of the project Matura work, e-marking of student exams papers, and promotion of e-testing formats in the system of external evaluation. She is a member of the Steering Committee of the FLIP+ e-assessment association.

Ben Rockliffe joined AlphaPlus in May 2021 as Deputy Director of Assessment. In his role, he provides technical input into the design and development of assessment services for AlphaPlus customers. His current significant projects include onscreen adaptive assessments for the Welsh and Scottish Governments.

Ben has worked in several different roles across the education and awarding sector. He worked on the Senior Management Team of two national awarding bodies where he led on qualification and assessment development, expanding the business into new markets and regulatory compliance. He has also worked in education policy at the Department for Education and the Training and Development Agency for Schools, where he led projects on workforce development, data improvement and legislation. Ben graduated in translation and interpretation in French and Spanish at the University of Bradford.

Bente Walgermo is associate professor at National Centre for Reading Education and Research in Noway. She specializes in research on reading and writing motivation, and assessment. She is leading large research- and development projects where the aim is to develop motivating assessments.

Bryan Maddox is Executive Director of Assessment MicroAnalytics Ltd; Director for Digital Assessment Futures at the Digital Education Futures Initiative, Hughes Hall, University of Cambridge; and Professor II (visiting professor) at the Centre for Educational Measurement (CEMO) at the University of Oslo. His research focuses on inclusion, diversity and user experience in digital educational assessments.

Carlos Palacios has a PhD in Language at the Federal University of Rio de Janeiro (UFRJ) and conducted post-doctoral research on public policies for literacy at the Department of Education at the Pontifical Catholic University of Rio de Janeiro (PUC-Rio). At CAEd/UFJF, he leads a team focused on research, innovation and partnerships with international institutions in large-scale assessment.

David West

Elodie Persem is a former teacher of literature specializing in working with disabled students. She has trained teachers and trainers for many years on the issues of school inclusion, managing heterogeneity and teaching cross-curricular skills. Since joining DEPP (Directorate for Evaluation, French Ministry of Education), she has been put in charge of the “AIR” unit (Accessibility, Innovation and Research) as a specialist in inclusive

education. This department was created to improve the accessibility of assessments and to develop partnerships with researchers in order to build assessments around the so-called 21st skills.

Eva de Schipper focuses on using assessment related data in innovative ways to create value for learners and teachers, through research as well as through the development of prototypes. In recent years, her work has focused on automated scoring of open ended questions, using log data from digital assessment to distill student strategies, and the use of recommender systems for personalized practice exams.

Franck Petrucci began his career in 2002 as a statistician at DEPP (Ministry de National Education, France). Since 2005, he has been a researcher at the Service de la recherche en éducation (SRED) of the Canton of Geneva (Switzerland). In 2018, he joined the team that conducts the large-scale assessments of PISA and UGK in Switzerland. His research focuses on measuring the contextual effects on students' achievement. Since 2020, he has been working on a doctoral thesis at the University of Grenoble Alpes (France) on teaching effectiveness in segregated education systems.

Dr. Jesse Dvorchak is a seasoned education professional with over 20 years of experience in educational leadership and innovation. She is an expert in assessment design, development, research methods, data analysis, and product implementation. Currently, Dr. Dvorchak serves as the Deputy-Director, Digital Products & Services at Cambridge University Press & Assessment, where she leads the ideation, development, and implementation of the organization's digital assessment products for over 10,000 schools in 160 countries around the world.

Jairo de Souza is an associate professor at the Federal University of Juiz de Fora, Brazil, and head of research at LApIC (Laboratory of Innovative Computational Applications) where he coordinates research on AI in Education, such as the automatic assessment of reading fluency in partnership with CAEd.

Dr. Joost Kruis received his Ph.D. in Item Response Mechanics from the Department of Psychological Methods at the University of Amsterdam. As a psychometric researcher he heads the AI research programme at Cito, where his current research focuses on applications of ML and AI in education. Additionally, he is involved in the norming and equating of the annual Dutch national secondary school exams.

Kevin Steinman is a Senior adviser at the Norwegian Directorate for Education and Training. As subject coordinator for upper secondary English, he has since November 2020 overseen the digitization process in summative high stakes assessments. This includes the introduction of listening items, with parallel accommodated video items for students with hearing challenges. Before becoming a bureaucrat, he worked for six years as a lecturer in English language literature and area studies at the University level. Kevin's first job after training as an English teacher was to become an a cappella singer.

After graduation he briefly took up a post in question paper preparation at Cambridge University Press & Assessment, but swiftly joined the Research and Evaluation division where he worked on, and in due course led, many studies in areas such as inter-board comparability, standards over time, equity and qualification evaluation. He subsequently worked extensively on the development and introduction of on-screen marking and on operationalizing data analytics. With over twenty years' experience in technical, research, and leadership roles at Cambridge University Press & Assessment, he now leads a programme of work in innovation and development.

Lucy Howarth

Luis Pereira dos Santos is currently the President of the Governing Board at the Institute for Educational Assessment (IAVE), responsible for national exams, standardized low stakes assessments, and international studies in Portugal. He represents Portugal in the PISA Governing Board (OECD), the General Assembly of the IEA, and chairs the Working Party of PISA-VET (OECD). Previously, he presided over the national examinations committee from 2011 to 2019. From 2005 to 2011, Luís was head of division for teacher training and director for multimedia educational resources at the Ministry of Education. He holds a degree in Physics Teaching and a Master's degree in Didactics for Science from the University of Lisbon.

Per Henning Uppstad is Professor of special needs education at the Norwegian Reading Centre, University of Stavanger. He holds a PhD in general linguistics, with research interests in technology enhanced learning, reading and writing, educational assessment and early literacy intervention.

Rui Pires is the Teacher Training Director at the Institute for Educational Assessment (IAVE) since 2019, where he also coordinates the dematerialization process of external assessments in Portugal. He has collaborated with IAVE since 2014. Previously, Rui was a Computer Science Teacher in secondary education since 2001 and a member of the governing board at Rainha Dona Leonor Secondary School in Lisbon for four years. He holds a Master's degree in Educational Sciences with a specialization in Educational Computer Science, a specialization in school management and administration, and a degree in Computer Science Teaching.

Sanjay Mistry Sanjay Mistry is Head of Research for Digital Assessment and Evaluation at Cambridge University Press and Assessment, International Education. The role involves leading on the definition of the research strategy that is pivotal to underpinning the organisation's strategic roadmap for the development of various digital assessment products and services that serve the international and UK markets, and working with schools, school groups and Associates to bring them on Cambridge's digital journey. Sanjay had a successful career in Primary education and leadership for 11 years before joining Cambridge International in 2012.

Sarah Hughes leads Research and Thought Leadership in the Digital High Stakes team in Cambridge University Press & Assessment. Her particular research interests are in digital assessment with an emphasis on validity. She currently carries out and applies research to support assessment practitioners developing and quality assuring digital assessments. She is passionate about applying research to influence decision making and policy and is building expertise in providing timely and relevant research evidence to practitioners in agile working environments.

Thierry Rocher is the Director of Student Evaluation at the Office for Student Assessment (DEPP) at the Ministry of Education in France. He is also the Chair of the IEA and the President and Founding member of the FLIP+ e-assessment association.

Tjeerd Hans Terpstra is an IT architect and technology advisor at Cito in the Netherlands. He works on various topics at Cito's R&D department, including computer adaptive testing and automated test assembly. He is a leader of the AI program at Cito.

Ulf Kroehne is a post-doc researcher at DIPF | Leibniz Institute for Research and Information in Education. His research interests include causal inference, psychometrics, technology-based assessment, adaptive testing, and log and process data.

Vali Huseyn is an educational assessment expert, with significant contribution to modernizing assessment practices across government and private sectors. At The State Examination Centre of Azerbaijan, he was instrumental in enhancing assessments and leading initiatives such as developing a unified testing platform and promoting assessment literacy. At Vretta, Vali leads strategic assessment initiatives, aligning them with organizational goals and educational standards, and drives innovative partnerships to boost assessment effectiveness and educational results.

Wagner Silveira Rezende has a PhD in Social Sciences and Education at the Federal University of Juiz de Fora (UFJF). Currently, he is a professor at UFJF Law School and general coordinator of research, assessment and technology projects at CAEd/UFJF.