GenAI for New Practices and Equity in Education (**GAINEd**)

1 Excellence

1.1 State of the art, knowledge needs and project objectives

Generative artificial intelligence (GenAI) offers great potential for the field of education, creating new opportunities to enhance learning, teaching, and education management. Given how recent this technological breakthrough is and how rapidly it is evolving, our understanding of its impacts on the education sector remains extremely limited. GenAI can personalize learning experiences, generate dynamic teaching content, automate administrative tasks, making education more efficient, engaging, and accessible. However, despite its promising potential, GenAI brings both opportunities and challenges. A critical yet underexplored question is how it impacts **equity in educational access and outcomes** and how it may reshape **pedagogy, teaching methods, and governance structures**.

The GAINEd project seeks to bridge this knowledge gap by uncovering the broader implications of GenAI on the equitable distribution of knowledge, the evolution of instructional practices, and the organization of educational systems for the future. While our scope spans different educational levels and stakeholders, the project is unified by a clear research focus: understanding the mechanisms through which GenAI affects student learning, teacher practices, and school leadership to develop evidence-based strategies for effective and equitable implementation. A close collaboration among our interdisciplinary team, spanning the fields of pedagogy, education economics, psychology, and leadership, allows us to to tackle GenAI's impact across multiple dimensions. To maximize impact, we will translate our findings into research-based best-practice recommendations for educational sector management and organizations.

We focus the proposal and state-of-the-art summary on three central themes. First, we examine **learning and teaching under the new AI paradigm**, including potential sources of inequality, such as gender disparities (Carvajal, Franco, & Isaksson, 2024). Second, we focus on competence development through the creation of a **pedagogical tool** that will help educators integrate and effectively utilize GenAI in their teaching. Third, we investigate **governance, organization, and management practices** at various educational levels to inform policies shaping the future of education in the AI era. The insights from GAINEd will equip stakeholders with the knowledge and tools to navigate challenges and harness the opportunities presented by GenAI.

According to state of the art research, GenAI increases productivity (Noy & Zhang, 2023; Brynjolfsson et al., 2025; Peng et al., 2023; Dell'Acqua et al., 2023) and seems to have inequality-reducing effects (Capraro et al., 2024) in the labor market. However, students that still are developing foundational learning skills may not benefit in the same way. As highlighted by a meta-study (Bond et al., 2024), evidence on **GenAI's impact on student learning** is scarce, despite a broad literature on effective AI applications in education. The few studies tackling the question on learning impacts through causal approaches as we propose in GAINEd have found mixed results. Randomized controlled trials (RCTs) show that students solving practice questions with ChatGPT perform worse when it is later removed in an exam, except if direct answers are restricted (Bastani et al., 2024). Others find negative effects on course engagement, but positive effects on exam performance for those who decide to take the optional course final exam (Nie et al., 2024). The emerging consensus: GenAI tools hinder learning for students who rely on them as a "crutch" but benefit those who seek explanations, as confirmed by the lab experiments in Lehmann et al. (2024).

Previous studies face criticism for design flaws (e.g., effectively preventing students in the control group from using GenAI) (Tan & Rajaratnam, 2024), or small sample sizes. A study that overcomes these issues is by GAINEd work package (WP) leader Franco. Franco, Irmert, and Isaksson (2025) invited 600 students to learn a new topic (Esperanto) in a lab, randomizing them into ChatGPT-assisted or Google Search-based learning. Despite concerns that GenAI harms learning, the study found no overall effect on performance or self-assessment. However, heterogeneous effects emerged, suggesting that female students may benefit, while low-GPA students may not.

Prior work offers little insight into which specific skills GenAI enhances or displaces, knowledge that is crucial for understanding its broader impact on learning and what tasks can be delegated without compromising educational outcomes. We also know little about how GenAI affects existing inequalities. If more advanced paid GenAI subscriptions or AI-based tutors significantly enhance student learning and lower-income students cannot afford them (Luckin, 2017), educational disparities may persist or worsen. Gender inequality may also worsen, as WP leader Franco finds that female students are less willing to pay for subscriptions (Carvajal, Franco, & Isaksson, 2024).GAINEd aims to bridge these gaps through innovative research designs that leverage diverse data types and contexts.

In terms of **teaching methods**, many see GenAI in education as an opportunity to revolutionize existing pedagogical practices (Perkins et al., 2024). A meta-systemic review found that AI facilitates personalized learning, constituting almost 40% of the identified advantages in the reviewed studies. Yet, educators remain cautious, with adoption largely driven by individual efforts or commercial AI tutors (Extance, 2023). This cautious approach stems from various factors, including ethical considerations, curriculum development, infrastructure limitations, and a lack of technical literacy (Bond et al., 2024). Aligning GenAI with pedagogy is a key challenge, as are issues of academic integrity, ethical use, and critical thinking. To date, the ethical and pedagogical implications of integrating GenAI into educational assessments remain under-explored in academic literature (Perkins et al., 2024).

The closest study to what we propose in GAINEd is by Perkins et al. (2024), who present a light-traffic system assessment scale to help educators adapt assessments to GenAI tools and provide clear guidelines on when and how students should use them while upholding academic integrity. While offering adaptable examples, the framework falls short of addressing curriculum development and teachers' technical literacy. Moreover, the tool remains untested for efficacy, logistical challenges, or user feedback. It is therefore crucial to develop a pedagogical tool that addresses these challenges and to evaluate its impact on student learning and teacher experiences. GAINEd aims to develop and test such a tool, ensuring equitable access for students and teachers from diverse backgrounds as a core consideration in its development and integration.

Regarding **governance structures**, so far, student adoption of GenAI tools has been largely unstructured, reflecting ongoing debates in education regarding readiness, ethics, trust, impact, and AI's value in learning. There is an urgent need for governance, regulation, research, and training to keep pace with the speed and scale of AI's transformation in education (Bond et al., 2024), as well as balanced regulation in its application (Dempere et al., 2023). Existing research underscores that integrating GenAI and other digital tools into assessments and learning environments requires more than a simplistic "allow/don't allow" framework. Instead, it is crucial to consider the complexities of digital literacy, ethical considerations, and collaborative efforts to establish academic integrity guidelines, with overly rigid or negative messaging about academic honesty potentially being counterproductive (Perkins et al., 2024).

There have been some efforts to introduce regulation, although not specifically in the education sector. Governments are implementing measures, with Europe leading the way by enacting the first AI law. However, efforts to develop specific frameworks and guidelines for educational institutions remain in their early stages. It is essential that these initiatives are grounded in rigorous research and strong conceptual foundations (Bond et al., 2024), and there is a clear need for research examining how GenAI influences governance, leadership, and organizational structures in education. GAINEd seeks to fill this gap by providing the first exploration of how educational leaders experience, use, and reflect on AI governance, AI-enhanced leadership, and AI-driven organizational efforts.

The GAINEd project has four primary objectives relating to different aspects of the education sector (described in detail in Section 1.2). **(WP1) Student learning.** We examine how students acquire different learning skills when GenAI is available and explores its impact on academic outcomes, such as grades. **(WP2) Pedagogical tool development.** We aim to develop and test a widely accessible, scalable pedagogical tool that optimizes AI use in the classroom to enhance student learning. **(WP3) Optimal teaching with AI.** Using the pedagogical tool developed in WP2, we assess its causal impact by comparing student performance and teacher satisfaction in courses where teachers have access to the tool vs. those where they do not. **(WP4) AI as a management challenge and opportunity.** We leverage unique access to school directors across education levels through two national leadership programs mandated by the Norwegian Directorate for Education (*Den nasjonale rektorutdanningen*). By examining the school

leaders' governance strategies, experiences, and organizational approaches to AI, we will provide crucial insights to inform educational authorities and stakeholders on best practices and the broader impact of AI on school leadership.

By meeting these project objectives and leveraging our team's diverse expertise and wide range of disciplines, the central contribution of the GAINEd project will be to establish a foundational understanding of how GenAI adoption influences individual academic outcomes and educational decision-making by students, teachers and school leaders. The project combines pedagogical competence on educational tools and learning and competence on experiment design and sophisticated survey and register data analysis to isolate and identify the impact of GenAI tools and the mechanisms through which they work. Furthermore, the project includes management, governance and organizational competence as well as experiences and feedback from several cohorts of school leaders, to develop best-practice recommendations for implementation of GenAI tools in the education sector.

A key focus across all work packages is understanding how GenAI impacts students, teachers, and school leaders across genders and socio-economic backgrounds. These insights will be crucial for designing targeted policy recommendations that promote equitable access and benefits. The focus on studying inequalities derived from GenAI access or use will also contribute toward achieving Sustainable Development Goal 10 on reducing inequality within and among countries. GAINEd insights will be relevant to researchers, policymakers, educators, and students, with strong potential for both academic and societal impact.

1.2 Research questions and hypotheses, theoretical approach and methodology

Work Packages

WP1. Student learning and grades (leader: Catalina Franco)

The main research question in WP1 is whether and how GenAI can enhance student learning and academic results. GenAI tools have the potential to both enhance and hinder learning processes, a duality rooted in the characteristic that they can generate human-like content. On the one hand, students using GenAI tools can harness them as powerful educational aids by adapting the tools to their individual learning paces and styles, using these tools more like a personalized, always-available tutor. On the other hand, there is concern that reliance on AI tools may diminish students' intrinsic motivation to learn and develop critical thinking skills. Easy access to answers can lead to surface-level understanding rather than deep learning, as students might bypass the cognitive processes required for problem-solving and comprehension. Grades and other academic results and decisions can also be affected as GenAI tools may also change student motivation and engagement. Ultimately, the impacts of GenAI hinge on whether it is used as a substitute for or a complement to student effort.

The challenge in establishing the causal effect of using GenAI tools on student learning and grades is threefold. First, GenAI use is difficult to observe and detection tools are not infallible (Sadasivan et al., 2023). Second, observational data from student assessments before and after the launch of GenAI tools would not reveal the true impact of GenAI on learning due to differences in GenAI adoption rates among students, potential spillover effects from classroom peers who are more skilled or frequent users of AI tools, and diverse teacher responses to the advent of this technology. Third, even randomized field experiments would not uncover the causal effect of GenAI use because it is practically impossible to prevent students in the control group to use GenAI tools.

Given these practical and methodological challenges, the GAINEd project proposes varied methodological approaches to provide evidence on the causal effects of GenAI on learning and grades. We focus on two main angles: (1) effects of GenAI use and exposure on individual learning and grades, and (2) spillover effects.

We propose laboratory experiments that will deepen our understanding of what skills are enhanced or crowded out by GenAI. For example, AI may facilitate analytical tasks such as problem-solving, while potentially weakening critical thinking or memory retention skills due to over-reliance. The experiments will have a teaching phase and a problem-solving phase where students are asked to learn and solve puzzles that measure cognitive executive functions. Through these experiments, we aim to explore foundational learning skills beyond a specific educational level, though participants are typically college students. In

the treatment condition, students will learn by interacting with an AI chatbot, while control students learn through traditional tutorials. In the problem-solving stage, students will not have access to any external aids, allowing us to measure whether students in both conditions can recall the sequence of steps and their efficiency in solving new puzzles. This design tests GenAI's impact on learning and problem-solving under controlled conditions, revealing whether it enhances understanding and retention or merely facilitates shortcut learning without deeper cognitive processing.

[hb!] **Research questions WP1**

1. What skills in the learning process does GenAI enhance or displace?
2. How do varying degrees of exposure to GenAI impact grades and academic choices?
3. Are there spillover effects from students who use GenAI on their peers who do not?

A second experiment will assess how students transfer AI-acquired knowledge to others. Specifically, we will compare students who learned with GenAI (treatment group) and those who used traditional methods (control group) in their ability to teach a peer. This setting provides insights into whether AI-based learning strengthens or weakens knowledge-sharing and collaborative learning outcomes.

We complement our experiments with a quantitative analysis using administrative data from Norway and Uruguay to assess how GenAI exposure influences academic outcomes. In Norway, we conduct a cohort study, comparing students with varying GenAI exposure during secondary and high school to pre-ChatGPT cohorts (before November 2022). While individual GenAI use is unobservable, we estimate its impact on grades and academic choices employing advanced machine learning methods to predict academic outcomes of exposed and non-exposed students who share the same student and family characteristics. The difference between the actual and predicted outcomes across cohorts will uncover the impact of GenAI exposure.

In Uruguay, we will leverage the collaboration with Ceibal, the agency in charge of integrating digital technologies into education in the country and our research partner in this proposal. To the best of our knowledge, Uruguay is the only country in the world that has access to students' web search histories—including GenAI chatbot usage—due to the government program providing a laptop to every student aged 6-15. By linking class-level data on GenAI use to academic records, we will test whether increased adoption rates within a classroom affect overall academic performance.

Finally, we recognize that GenAI may influence how instructors teach. To account for this in the analyses, we will conduct surveys with instructors in both countries to assess the extent to which they have integrated GenAI into their teaching methods, evaluations, and preparation.

Building on the WP leader's previous work, we will examine how gender and the classroom gender composition influences outcomes. By identifying student characteristics and GenAI usage patterns linked to stronger learning outcomes, this WP will contribute to research on technology in education, the economics of education, and educational psychology. The findings will also offer valuable guidance for educators, students, and institutions on effectively leveraging GenAI tools to enhance learning, with both short- and long-term benefits. The results from WP1 will inform and complement the development of the pedagogical tool in WP2. This interrelation leverages the WP leaders' expertise in quantitative and pedagogical methods, respectively, presenting a well-rounded approach.

WP2. Pedagogical tool development (leader: Crina Damşa)

This work package focuses on developing a flexible and adaptable pedagogical tool that structures AI interactions to support student learning. It addresses the need to define appropriate applications of GenAI in education (Perkins & Roe, 2024b), incorporate GenAI into curriculum design (Bahroun et al., 2023), establish clear guidelines for its responsible use (Cotton et al., 2024), and ensure transparency in its application (Perkins & Roe, 2024a).

The tool will integrate key elements of custom AI tutors so that students engage in critical thinking rather than passively receiving answers. It will provide strategies for personalized feedback and self-paced learning, allowing students to tailor AI interactions to their individual needs. Additionally, the tool

will emphasize inquiry-based learning by encouraging students to ask questions, explore concepts in depth, and actively engage in problem-solving. Another essential feature will be training students to critically evaluate AI-generated responses, assessing their accuracy, relevance, and potential biases.

Designed to be customizable, the tool will offer teachers structured guidance, in the form of a manual, on effectively incorporating GenAI into course assignments. The most basic version will consist of adaptable assignment templates that clearly demonstrate how AI can support concept development. These templates will include specific guidelines on student-AI interactions, outline submission requirements—such as logs of GenAI chatbot conversations alongside student responses—and provide clear grading criteria to ensure transparency in assessment.

Research questions WP2

1. What design principles ensure that a pedagogical tool effectively integrates GenAI into teaching and learning while maintaining academic integrity?
2. What strategies and support mechanisms best enable teachers to develop the knowledge and skills needed to integrate GenAI into their teaching?

The specific contents of the tool will include: (1) A set of example course assignments customizable for courses in different fields. (2) Clear guidelines on what students need to submit and a suggestion on how to grade it. (3) Language to include in course plans to introduce the assignments that require the use of GenAI. (4) Guidelines on ethics and academic integrity that teachers can rely on to discuss ethical issues in the classroom, emphasizing using a tool as a complement rather than asking for direct answers, as well as the critical evaluation of GenAI output. (5) A tutorial for teachers so they can try solving the assignments themselves and overcome the technical literacy barrier. The tool will be developed following pedagogical principles of interactive learning environments and the contents will benefit from discussion from teachers and students to share insights on best practices, identify challenges, and develop effective strategies to improve the tool and pedagogical strategies. It will also incorporate the insights gained from WP1 and be piloted and tested to ensure a smooth implementation at scale. The results and feedback will be used to provide recommendations that allow for future tool development, beyond the duration of the project.

Leveraging the WP leader's expertise, this work package will move the frontier in integration of digital technologies and pedagogical innovation in the AI era. By ensuring the tool is both accessible and flexible, we will bridge the gap between custom-made AI tutors and commercially available AI chatbots with enormous impact potential due to its scalability and adaptability across different fields and contexts. WP3 will rigorously assess the tool's effectiveness, leveraging the complementary strengths of pedagogical and quantitative approaches from our team members.

WP3. Optimal teaching with AI (co-leaders: Astrid Oline Ervik and Kjell Salvanes)

The central research question in WP3 is whether providing teachers with a pedagogical tool, as outlined in WP2, positively impacts students' learning outcomes and experiences, as well as teachers' motivation and performance. Teachers are central to the AI revolution, yet GenAI integration faces significant hurdles. WP3 aims to address these challenges by putting the pedagogical tool from WP2 to the test.

Identifying the impact of teachers' GenAI use in their courses presents two key challenges. First, there is no systematic data on whether and to what extent teachers integrate GenAI into student learning, as they may use it mainly for course planning or grading. Second, even if we measure GenAI adoption, its use may correlate with teacher quality. If higher-quality teachers are more likely to adopt GenAI and their students perform better, we risk misattributing improved outcomes to GenAI rather than teacher effectiveness. Addressing these biases is crucial for understanding GenAI's true effect on learning.

To isolate GenAI's impact as a teaching tool, we will conduct two RCTs across different educational levels in Norway and Uruguay, capturing effects in diverse contexts. Randomly selected teachers will receive access to the WP2 pedagogical tool but can choose whether and how much to use it. They will report their usage, allowing us to measure adoption rates and gain insights into what works and what does not in real-world implementation.

Research questions WP3

1. How does incorporating a pedagogical tool designed to harness GenAI impact students' academic outcomes and learning experiences?
2. Do these effects vary by baseline academic skills and gender?
3. How do teachers experience incorporating this tool in their courses? Do they perceive it as beneficial for student learning?

In Norway, we will survey university professors on their GenAI usage and perceptions of its impact on teaching and learning. From the respondents, a random subset will be assigned to the treatment group and given access to the pedagogical tool, with instructions to not share it during the trial. Control group professors will continue with their usual course plans. At the end of the semester end and for a few more periods, we will collect student outcomes such as completion of their study program and enrollment into further education from administrative data. We will conduct follow-up surveys with professors in both groups to assess their experience with their tool and feedback for improvement. Norway provides an ideal setting due to high, but uneven GenAI adoption across study fields (Ministry of Education, 2024), allowing us to assess whether the tool benefits students equally across different subjects. Additionally, administrative data enables us to analyze individual student outcomes and overall effects of integrating GenAI in terms of equity and diversity in education.

In Uruguay, we will collaborate with Ceibal to conduct a complementary study with teachers of students aged 10 to 15. Through Ceibal we will have access to all schools in the public education system and we will have access to student grades, both of which are infeasible in Norway. Another difference with the study in Norway is that we will collect a series of survey measures of beliefs, attitudes and non-cognitive skills that may be associated with effective learning. Survey measures along with academic and demographic background will allow us to identify mechanisms through which GenAI tools aid or hinder learning. One important advantage of doing this study in a developing country is that access to cutting-edge technologies such as GenAI is more limited than in developed countries (Mannuru et al., 2023), allowing us to study first adoption and the learning curve.

The insights gained from the RCTs with teachers will offer the first large-scale evidence on integrating AI into teaching, while providing a scalable, accessible alternative to highly curated AI tutoring out of reach for most educators and students. These findings will push the research frontier in both education and labor economics, and will be informative to school leaders' reflections on governance and managerial practices covered in WP4.

WP4. AI as a management challenge and opportunity (leader: Synnøve Nesse)

The main research question guiding WP4 is how AI may enhance education management. As AI reshapes teaching and learning, its role in governance, leadership, and organization is increasingly important (Fullan, Azorín, Harris, & Jones, 2024). However, its use in educational management remains largely ungoverned, unstructured, and highly variable across contexts (Tuomi et al., 2022; Fullan et al., 2024). Understanding AI's influence across all educational levels—from early childhood to higher education—will help policymakers adopt it strategically while mitigating risks such as inequality instead of leaving it to chance. In this nascent field, explorative research is needed to examine how educational leaders experience, use, and reflect on AI governance, AI-enhanced leadership, and AI-organizing efforts, which is the focus of this WP.

First, educational management is governed by national authorities through laws and policies that should include AI governance. However, AI technologies evolve rapidly, making policy development challenging (Fullan et al., 2024). For example, Norwegian policies mention AI in learning and teaching, but not in management (Meld. St. 34, 2023-2024). Further, educational leadership, while seemingly hierarchical, is in practice often distributed, team-based, individual and autonomous (Spillane et al., 2001), complicating AI policy implementation. Finally, management may not be aware of GenAI tools' use or effects. This may create differences in management practice depending on gender (Carvajal et al., 2024),

geography, level and school size (Isaksen et al., 2024). As a result, AI governance inconsistency may produce inequalities in the absence of clear policies and practices.

Second, AI shows promise in enhancing leadership effectiveness, and educational leaders are experimenting using it (Adams & Thompson, 2025). AI can support key leadership functions such as decision-making and task execution, managing relations and conflicts, and achieving creativity and innovation (Yukl, 2012). Studies suggest AI improves decision-making and accelerates task completion in both administration and teaching (Adams & Thompson, 2025). As AI can aid development of empathy and perspective-taking (Langley et al., 2022), it may help manage stress and conflict, and support building relational holding environments. Additionally, AI can foster creativity by shifting leaders towards more innovative work (Kesim, Atmaca, & Turan, 2025). However, leaders vary in AI experimentation, making it crucial to examine these differences and their impact on inequalities, to close the gap, across educational levels and contexts.

A third aspect of educational leadership concerns how using AI is organized in different management structures. Compared with implementing prior digital technologies, AI may be used by a few leaders who form a group willing to experiment with the technology, and then share it with others. Another version is that leaders begin using it in unstructured ways in their individual practice (Rukadikar & Khandelwal, 2024). However, both approaches are suboptimal when it comes to harnessing experience, shared learning and deeper exploration of risks and ethics, and achieving organizational development, as transfer to others may prove difficult (Kesim et al., 2025). To date, knowledge about how educational leadership organizes efforts to integrate AI into their leadership is limited, which gives rise to our three research questions.

Research questions WP4

1. How do leaders experience (the absence of) governance (e.g. policy documents) regarding AI in educational management?
2. How do educational leaders experiment with AI in their core leadership tasks?
3. How do leaders organize their understanding, use, and experience of AI within their institutions?

To answer our research questions, we will conduct a longitudinal mixed-methods study with educational leaders in two national leadership programs (one for kindergarten and one for primary through upper secondary school leaders) mandated by the Norwegian Directorate for Education and Training and run by the Norwegian School of Economics (NHH) and Administrativt forskningsfond (AFF). The unique government-backed program ensures consistency and high-quality leadership training across Norway. During these programs, leaders meet ten times over two years, allowing us to collect baseline data and track changes over time for six cohorts (N=300).

For our first study, examining all three research questions, we will interview leaders (in groups of 5-7) to explore how their experiences and reflections on AI change over time. We will also give them access to the pedagogical tool developed and tested in WP2 and WP3 to receive feedback from the school leader perspective. Second, we will conduct an action-research study using a "nudging intervention" (O'Meara et al., 2022) with experiment and control groups, encouraging GenAI use in leadership tasks alongside critical reflection and learning. For our third study, we will provide experience-based policy recommendations for national authorities, school owners, principals, and kindergarten directors regarding a) governing AI adoption in educational management, b) experimenting with AI in leadership, and c) structuring AI integration to maximize learning, enhance benefits, and reduce inequalities.

Insights from WP4 will advance research on leadership effectiveness and educational leadership while fostering strong engagement with educational leaders on the findings and methods from WPs 1–3.

1.3 Ethical issues

The GAINEd project will ensure that scientific methods are applied in a professionally sound manner, and that researcher's integrity and freedom is secured. The research team does not have any strings or

obligations attached to stakeholders such as the ed-tech industry or government institutions. To ensure accountability and reproducibility we will openly publish datasets and research results from all WPs, while at the same time ensuring compliance with privacy requirements.

The Centre for Applied Research (SNF), where the project will be hosted, follows the Guidelines for Research Ethics in the Social Sciences and the Humanities issued by the National Committee for Research Ethics in the Social Sciences and the Humanities (NESH). NHH and SNF share a joint Research Integrity Committee and an Institutional Review Board (IRB).

Ethical approval will be obtained for all experiments involving students and teachers, and we will strictly adhere to established ethical protocols to protect participant rights, privacy, and well-being. The Norwegian Agency for Shared Services in Education and Research (Sikt) serves as SNF's Data Protection Officer. All projects involving the processing of personal data are reported to Sikt. We will obtain informed consent from all participants, clearly explaining the study's purpose, their rights, and how their data will be used. Personal and sensitive data will be stored at NHH Safe, a server used by NHH and SNF for safe storage of research data.

The project will emphasize fair and mutually beneficial collaboration with Ceibal, the national agency in Uruguay. Continuous ethical reflection and discussion will ensure that cultural and contextual factors are considered in research design.

Project risks

We consider three types of risks for the successful completion of GAINEd. First, large-scale RCTs entail diverse risks such as non-compliance (e.g., treatment teachers not using the pedagogical tool or control teachers adopting it) and ethical concerns due to unsupervised use of GenAI in control groups. We consider these risks manageable. We will use standard econometric techniques for compliance issues (Angrist & Imbens, 1995) to ensure we recover the causal effects of using the pedagogical tool. In addition, any unsupervised use of AI tools reflects the current status quo, which we take as a given for the purposes of the study since we cannot interfere with school policies or individual decisions regarding GenAI use.

Second, the broad focus on participants and target groups across different educational levels and countries poses a potential risk, as it requires significant capacity for data collection and analysis. To mitigate this risk, we will implement a structured, phased approach that prioritizes key research questions while ensuring feasibility. We will leverage existing administrative data and educational platforms in Norway and Uruguay to streamline data collection, collaborate with local partners to facilitate access to schools and universities.

Third, regarding data management and processing, the risks are minimal as we have expertise on data collection form lab experiments and RCTs, and have used administrative data from Norway and Uruguay as part of earlier and on-going projects.

1.4 Novelty and ambition

The project offers a highly innovative and comprehensive analysis of learning effects, new practices and equity in education. Frontier cutting-edge GenAI technology has the potential to redefine traditional educational practices, methods, and outcomes. We will use a wide array of methodologies and identification strategies, ambitiously designed to uncover the causal impacts of GenAI on education and to illuminate the key mechanisms driving these effects, while always also focusing on how to enhance equity.

The GAINEd project is driven by a team of five talented, highly ambitious, and prolific researchers, each bringing expertise in various education-related topics and methodologies. As far as we know, we are the first research team to integrate surveys with Norwegian administrative data to study GenAI in education, and among the first to design laboratory and field experiments with students and teachers specifically focused on the implications of GenAI. Our collaboration with Ceibal in Uruguay which, with their focus on incorporating technology in education, will also prove to be invaluable to gain insights from non-developed countries.

The feasibility of the project is supported by the team of researchers at the Centre of Excellence FAIR at NHH. This additional expertise will ensure that our research not only breaks new ground but also achieves its objectives with precision and impact.

2 Impact

2.1 Potential for academic impact of the research project

The GAINEd project is uniquely positioned to make significant academic contributions thanks to the diversity of strengths of the team members and varied methodological approaches, including the development of a new pedagogical tool. By employing diverse research designs—including laboratory experiments, RCTs, quasi-experimental and mixed methods— and creatively combining different data sources, we will build a foundational understanding of how GenAI adoption impacts academic outcomes and educational decision-making. This approach will not only shed light on GenAI's effects on teaching and learning, but also generate insight on the governance, management and organization dimensions in the education sector.

The project aims to publish at least two articles per work package in high-impact journals across economics, general science, leadership, and pedagogy. With a team of researchers with strong publication records, the GAINEd project is well-positioned to make a significant academic impact and contribute to the expanding interdisciplinary network studying GenAI's effects on education.

2.2 Potential for societal impact of the research project

As discussed above, the lack of coordinated policy regarding GenAI use presents significant risks. Without a clear understanding of GenAI's implications, educational institutions may fail to harness its potential benefits while also exposing themselves to serious challenges such as student unsupervised use of these tools affecting safety, equity, and learning (Ta & West, 2023). The GAINEd project will provide much-needed insights into how GenAI can be effectively integrated into education. It will help educational institutions develop guiding principles, create training resources for teacher professional development, and empower educators to make the best use of GenAI in the classroom. We aim to establish and strengthen relationships with educational institutions and government authorities in Norway and Uruguay, offering evidence-based recommendations to help them and institutions in other countries navigate the evolving landscape of GenAI in education.

GAINEd will help improve our understanding of societal challenges by providing:

- A better understanding of the benefits and pitfalls of GenAI in the classroom, which can inform schools, educational authorities, students and parents.
- A thorough understanding of the distributional effects of GenAI, which is of importance for a wide array of policy questions, including alleviating systematic obstacles for initially disadvantaged students and female students in technology-related fields.
- Informed recommendations for improving school and university practices and policies on GenAI use, reducing inconsistencies in student outcomes caused by unclear AI guidelines.

2.3 Measures for communication and exploitation

To reach a wider audience, we will pursue extensive press coverage, publish policy briefs, and give public lectures. We also plan to offer a course on AI in education through platforms like the Bergen Summer Research School. Dissemination through newspapers, podcasts, blogs, and active social media engagement will make our findings accessible to the public. We will target education authorities, school officials and students, for whom the research results are most relevant.

The gender differences in GenAI adoption paper by team member Franco has already gained visibility in major outlets such as The Economist, The Conversation, Khrono, and forskning.no. GAINEd will also leverage the dissemination platforms of the Center of Excellence FAIR (Research Council of Norway, 2017–2027) and SNF, which engage a broad audience through multiple social media channels.

We will organize a two-day international workshop, with one day dedicated to academics and the second open to stakeholders and policymakers at national and local levels in Norway. This includes universities, schools, the Ministry of Education and Research, and the Norwegian Directorates for Education and Training (Udir) and for Higher Education and Skills (HK-dir). The second day aims to disseminate findings and gather feedback to maximize the research's impact on non-academic stakeholders. Additionally, we will present at international conferences and deliver seminar talks worldwide to further engage with the academic community and policymakers.

3 Implementation

3.1 Project manager and project group

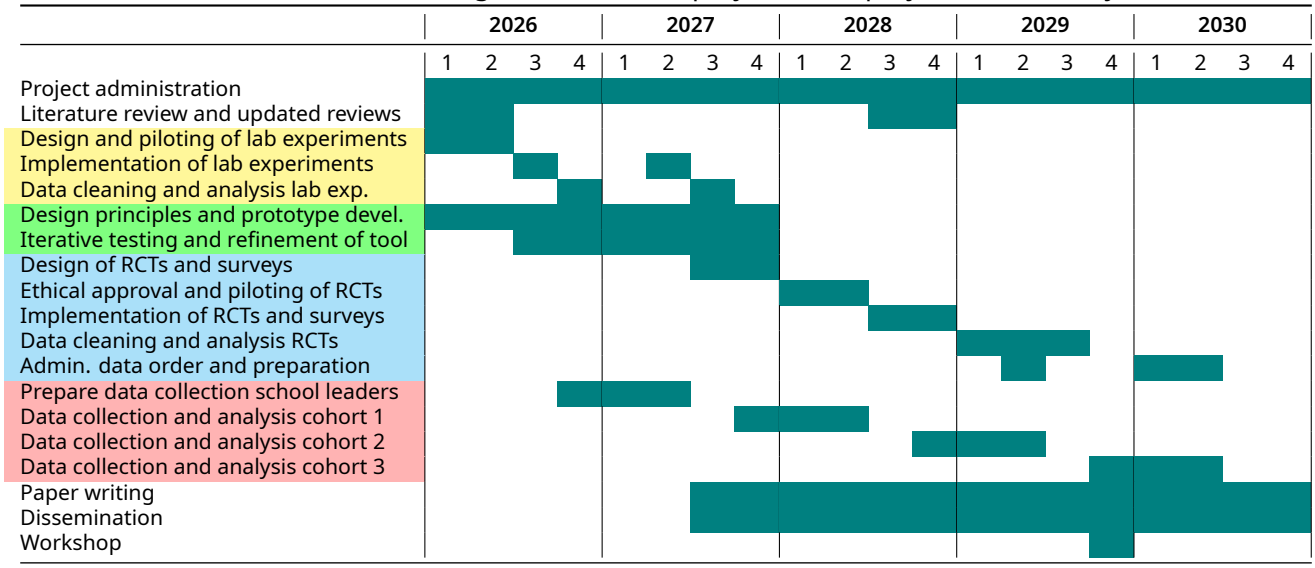
Principal investigator **Kjell Salvanes** is an Economics Professor at NHH and co-director of the Center of Excellence, FAIR. He is one of the top 5 percent of economists in the world and a leading expert on education and labor market research, and inequality. He has served on numerous national expert committees. He has been a project manager for a long list of large-scale research projects, most recently on the effects of robots on workers and households and on interventions to increase high school completion and upskilling or reskilling for low-skilled workers.

The main project team includes four additional researchers. **Catalina Franco** (economist, Ph.D. University of Michigan) and **Astrid Oline Ervik** (economist, Ph.D. University of Cambridge) are researchers at the Centre for Applied Research (SNF) at NHH and are affiliates at the center of excellence FAIR. Franco is actively developing an agenda on the impacts of GenAI on education and the labor market, with two papers directly relevant to this proposal. Ervik is a research group leader, conducts research on the education sector, and has served on an expert committee appointed by the Ministry of Education. Both researchers have extensive experience in managing and contributing to large research projects as well as designing and implementing RCTs and collecting student surveys. **Crina Damşa** (Educational Sciences, Ph.D. University of Oslo) is a professor at the University of Oslo Department of Education. Her work focuses on digital learning and pedagogical design practices, with previous experience developing both pedagogical and technological tools. **Synnøve Nesse** (psychologist, Ph.D. NHH) is a senior researcher and research group leader at SNF with extensive experience in project management, unique expertise in leadership during uncertainty, change management, and the use and scaling of AI.

3.2 Project organization and management

Salvanes will overseee the whole project and co-lead WP 3. Franco will lead WP1. Damşa will lead WP2. Ervik will co-lead WP3 with Salvanes. Nesse will lead WP4. University of Chicago professor Ariel Kalil (developmental psychologist) will contribute to WP1-3. NHH Professor in strategy and management Alexander Madsen Sandvik will participate in WP4. Researcher Marcela Gómez-Ruiz from Ceibal Uruguay will work on WP 1 and 3. All WP leaders will work in close collaboration to maximize the benefits of the interdisciplinary team. The work plan with the main activities is outlined in Table 1.

Table 1: Workflow and timing of the different project tasks (project start January, 2026)



Color coding: WP1 (yellow), WP2 (green), WP3 (blue), WP4 (red). No color means activities transversal to all WPs.

References

Adams, D., & Thompson, P. (2025). Transforming school leadership with artificial intelligence: Applications, implications, and future directions. *Leadership and Policy in Schools*, *24*(1), 77–89.

Angrist, J., & Imbens, G. (1995). *Identification and estimation of local average treatment effects.* National Bureau of Economic Research Cambridge, Mass., USA.

Bahroun, Z., Anane, C., Ahmed, V., & Zacca, A. (2023). Transforming education: A comprehensive review of genera-

tive artificial intelligence in educational settings through bibliometric and content analysis. *Sustainability*, *15*(17), 12983.

Bastani, H., Bastani, O., Sungu, A., Ge, H., Kabakcı, O., & Mariman, R. (2024). Generative ai can harm learning. *Available at SSRN*, *4895486*.

Bond, M., Khosravi, H., De Laat, M., Bergdahl, N., Negrea, V., Oxley, E., ... Siemens, G. (2024). A meta systematic review of artificial intelligence in higher education: A call for increased ethics, collaboration, and rigour. *International Journal of Educational Technology in Higher Education*, *21*(1), 4.

Brynjolfsson, E., Li, D., & Raymond, L. (2025). Generative ai at work. *The Quarterly Journal of Economics*.

Capraro, V., Lentsch, A., Acemoglu, D., Akgun, S., Akhmedova, A., Bilancini, E., ... others (2024). The impact of generative artificial intelligence on socioeconomic inequalities and policy making. *PNAS nexus*, *3*(6).

Carvajal, D., Franco, C., & Isaksson, S. (2024). Will Artificial Intelligence Get in the Way of Achieving Gender Equality? *NHH Dept. of Economics Discussion Paper*(03).

Cotton, D. R., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of chatgpt. *Innovations in education and teaching international*, *61*(2), 228–239.

Dell'Acqua, F., McFowland, E., Mollick, E. R., Lifshitz-Assaf, H., Kellogg, K., Rajendran, S., ... Lakhani, K. R. (2023). Navigating the jagged technological frontier: Field experimental evidence of the effects of AI on knowledge worker productivity and quality. *Harvard Business School Technology & Operations Mgt. Unit Working Paper*(24-013).

Dempere, J., Modugu, K., Hesham, A., & Ramasamy, L. K. (2023). The impact of chatgpt on higher education. In *Frontiers in education* (Vol. 8, p. 1206936).

Extance, A. (2023). Chatgpt has entered the classroom: how llms could transform education. *Nature*, *623*(7987), 474–477.

Franco, C., Irmert, N., & Isaksson, S. (2025). Does AI help or hurt student learning?

Fullan, M., Azorín, C., Harris, A., & Jones, M. (2024). Artificial intelligence and school leadership: challenges, opportunities and implications. *School Leadership & Management*, *44*(4), 339–346.

Isaksen, S., Kristin, N., Jenssen, P., & Eide, G. A. (2024). Transforming distance learning with artificial intelligence in remote areas of norway during the pandemic. *International Journal of Education, Science, Technology, and Engineering (IJESTE)*, *7*(1), 1–12.

Kesim, E., Atmaca, T., & Turan, S. (2025). Reshaping school cultures: Ai's influence on organizational dynamics and leadership behaviors. *Leadership and Policy in Schools*, *24*(1), 117–136.

Langley, C., Cirstea, B. I., Cuzzolin, F., & Sahakian, B. J. (2022). Theory of mind and preference learning at the interface of cognitive science, neuroscience, and ai: A review. *Frontiers in artificial intelligence*, *5*, 778852.

Lehmann, M., Cornelius, P. B., & Sting, F. J. (2024). Ai meets the classroom: When does chatgpt harm learning?

Luckin, R. (2017). Towards artificial intelligence-based assessment systems. *Nature Human Behaviour*, *1*(3), 0028.

Mannuru, N. R., Shahriar, S., Teel, Z. A., Wang, T., Lund, B. D., Tijani, S., ... others (2023). Artificial intelligence in developing countries: The impact of generative artificial intelligence (AI) technologies for development. *Information Development*.

Meld. St. 34. (2023-2024). *En mer praktisk skole. bedre læring, motivasjon og trivsel på 5.–10. trinn.* `https://www.regjeringen.no/contentassets/dcc936e85f4c43bf90ecf0bafabb00a8/no/pdfs/stm202320240034000dddpdfs.pdf`.

Ministry of Education. (2024). *2023 studiebarometeret.* `https://www.studiebarometeret.no/no/artikkel/5`.

Nie, A., Chandak, Y., Suzara, M., Malik, A., Woodrow, J., Peng, M., ... Piech, C. (2024). *The gpt surprise: Offering large language model chat in a massive coding class reduced engagement but increased adopters' exam performances* (Tech. Rep.). Center for Open Science.

Noy, S., & Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Science*.

O'Meara, K., Culpepper, D., Lennartz, C., & Braxton, J. (2022). Leveraging nudges to improve the academic workplace: Challenges and possibilities. *Higher Education: Handbook of Theory and Research: Volume 37*, 277–346.

Peng, S., Kalliamvakou, E., Cihon, P., & Demirer, M. (2023). The impact of AI on developer productivity: Evidence from GitHub copilot. *arXiv:2302.06590*.

Perkins, M., Furze, L., Roe, J., & MacVaugh, J. (2024). The artificial intelligence assessment scale (aias): A framework for ethical integration of generative ai in educational assessment. *Journal of University Teaching and Learning Practice*, *21*(6), 49–66.

Perkins, M., & Roe, J. (2024a). Academic publisher guidelines on ai usage: A chatgpt supported thematic analysis. *F1000Research*, *12*, 1398.

Perkins, M., & Roe, J. (2024b). Decoding academic integrity policies: A corpus linguistics investigation of ai and other technological threats. *Higher Education Policy*, *37*(3), 633–653.

Rukadikar, A., & Khandelwal, K. (2024). Leadership development through self-upskilling: role of generative artificial intelligence. *Development and Learning in Organizations: An International Journal*, *38*(4), 27–30.

Sadasivan, V. S., Kumar, A., Balasubramanian, S., Wang, W., & Feizi, S. (2023). Can ai-generated text be reliably detected? *arXiv preprint arXiv:2303.11156*.

Spillane, J. P., Halverson, R., & Diamond, J. B. (2001). Investigating school leadership practice: A distributed perspective. *Educational researcher*, *30*(3), 23–28.

Ta, R., & West, D. M. (2023). *Commentary: Should schools ban or integrate generative AI in the classroom?* `https://www.brookings.edu/articles/should-schools-ban-or-integrate-generative-ai-in-the-classroom/`.

Tan, S., & Rajaratnam, V. (2024). Critique of generative ai can harm learning study design. *Available at SSRN 4898213*.

Tuomi, I., Holmes, W., & Miller, R. (2022). Charting the futures of artificial intelligence in education. *European Journal of Education*, *57*(4).

Yukl, G. (2012). Effective leadership behavior: What we know and what questions need more attention. *Academy of Management perspectives*, *26*(4), 66–85.