leaders

RESPONSIBLE AI

SUCCESS STORIES

www.aileaders-project.eu

Co-funded by the European Union

Co-funded by the European Union. Views and opinions expressed are however those of the author or authors only and do not necessarily reflect those of the European Union or the Foundation for the Development of the Education System. Neither the European Union nor the entity providing the grant can be held responsible for them.

Success Stories

CLICK TO VIEW

www.aileaders-project.eu

Fairness

Balancing Fairness and Transparency in AI Fraud Detection

SETTING

Sérgio Jesus is a Research Data Scientist at Feedzai, a global company providing AI-driven solutions for financial fraud detection. With five years of experience at Feedzai, Sérgio works on developing and deploying AI systems that help financial institutions combat fraud.

Feedzai's AI solutions are used across various geographies, serving clients in North America, Europe, Australia, Asia, and South America. The core application of AI in Feedzai's systems is fraud detection, but the company also develops tools for anti-money laundering (AML).

Sé tems at Fee pri utomatically c transactions, flagging suspicious activities, and helping financial institutions prevent fraud in real time. This work has a profound ethical dimension, particularly when considering fairness and transparency.

> *"Models can still infer sensitive attributes from other data, which might lead to unintended bias."*

CHALLENGE

One of the most significant ethical challenges Feedzai faces is ensuring fairness in its AI systems. Sérgio draws an example from the financial industry, referencing cases such as the Apple Card incident, where algorithmic biases resulted in a credit card offering women lower credit limits than men, despite having identical financial profiles. While Feedzai does not control credit limits, its systems have the power to block accounts and transactions. The potential for unintended biases in these actions underscores the importance of fairness in their fraud detection systems.

aileaders-project.eu
leaders
FACULTY OF MANAGEMENT University of Lodz
FEP
SCHOOL OF ECONOMICS AND MANAGEMENT UNIVERSITY OF PORTO
celtech software innovations ltd
European E-Learning Institute
ACEEU ACCREDITATION COUNCIL FOR ENTREPRENEURIAL & ENGAGED UNIVERSITIES
upf. BARCELONA SCHOOL OF MANAGEMENT
Co-funded by the European Union
CLICK TO VIEW

To address fairness, Feedzai provides clients with tools for auditing models. These tools allow data scientists to assess whether decisions are biased toward particular groups, ensuring that fairness is maintained across sensitive attributes, such as race, gender, and nationality. Feedzai also includes fair model training features, enabling clients to balance performance metrics with fairness metrics during the model training process. This ensures that while the model's accuracy is optimised, fairness metrics, such as equal opportunity, are not sacrificed.

In addition to fairness, another key challenge is transparency and explainability. Feedzai's AI systems provide automated decisions to fraud analysts, who rely on explanations from the model to make more informed decisions. Sérgio notes that it's vital to provide clear, interpretable explanations to help analysts understand why a transaction was flagged as fraudulent. This approach improves decision-making speed and accuracy, while also ensuring accountability.

CONCLUSION

Feedzai has implemented several advanced tools and methodologies to tackle the challenges of bias and transparency in AI. One such tool is the FairGBM (Gradient Boosting Model), which Sérgio explains is designed to optimise both performance and fairness simultaneously. FairGBM ensures that fairness constraints, such as equal opportunity across different demographic groups, are respected at each boosting step of the model's training process. This allows the system to correct potential fairness violations while maintaining accuracy.

Additionally, Feedzai has contributed to Aequitas, an open-source tool designed for bias detection in AI models. Aequitas calculates fairness metrics, such as precision and recall, for each demographic group in the dataset. This allows Feedzai to monitor how well their models perform across different user segments, ensuring that no group is unfairly treated or disproportionately affected by fraud detection algorithms.

One of the significant challenges Sérgio mentions is the lack of access to sensitive data needed to detect discrimination. For example, many clients do not collect data on sensitive attributes such as race or gender, believing this prevents discrimination. However, Sérgio warns that models can still infer these attributes from other data, which might lead to unintended bias. Feedzai works to educate clients on the importance of measuring fairness metrics and advises them on how to collect the necessary data to ensure their AI systems remain fair and unbiased.

Finally, Feedzai addresses the challenge of complex fraud patterns with explainability tools. Fraud analysts often deal with adversarial and fast-evolving fraud techniques. In response, Feedzai adopts feature attribution methods like SHAP, LIME, etc. for model debugging keeping in mind the final users, allowing data scientists to interpret the model's behaviour and taking faster and better decisions. For analysts, Feedzai's tools provide more specific, rule-based explanations that reflect common fraud patterns, such as card cloning or phishing attacks. These explanations are essential for identifying evolving fraud schemes and improving the model's robustness over time.

DIVE DEEPER

FairGBM – A tool developed by Feedzai that applies fairness constraints during model training to balance accuracy with fairness across demographic groups.

Aequitas – An open-source bias detection tool that measures fairness metrics like precision and recall for different demographic groups, helping ensure equitable model performance.

Explanation Methods (such as SHAP & LIME) – Used for model debugging to provide transparency into AI decision-making processes and help analysts better understand why certain transactions were flagged as fraudulent.

Co-funded by the European Union

This project has been funded with support from the European Commission. The author is solely responsible for this publication (communication) and the Commission accepts no responsibility for any use may be made of the information contained therein. In compliance of the new GDPR framework, please note that the Partnership will only process your personal data in the sole interest and purpose of the project and without any prejudice to your rights.

TRANSPARENCY & EXPLAINABILITY

Building Trust in AI: Dynamic Pricing

SETTING

André Morim is a Senior Consultant at LTP Labs , a consulting firm dedicated to empowering decision-making through data analytics and AI-driven solutions. The company initially operated in Portugal but has since expanded internationally, delivering projects in Southeast Asia, Africa, and the Americas. LTP Labs specialises in providing AI solutions across multiple sectors, including retail and healthcare, to help organisations make data-driven decisions and optimise their operations.

One of the most notable projects André discussed involved working with a large Portuguese retailer on implementing a dynamic pricing system for perishable goods. This AI-powered model aimed to reduce food waste by adjusting the pricing of items approaching their expiration dates, while simultaneously maximising profit margins. The project involved modelling customer behaviour in response to discounts and balancing various factors like product availability, type, and discount percentage.

aileaders-project.eu
leaders
FACULTY OF MANAGEMENT University of Lodz
FEP
SCHOOL OF ECONOMICS AND MANAGEMENT UNIVERSITY OF PORTO
software innova
European E-Learning Institute
ACEEU
ACCREDITATION COUNCIL FOR ENTREPRENEURIAL & ENGAGED UNIVERSITIES
upf.
BARCE SCHOOL MANAGEMEN
Co-funded by the European Union
CLICK TO VIEW

> *“By involving stakeholders early in the development process, AI models are not only trusted but also optimised for practical use.”*

CHALLENGE

aileaders-project.eu
leaders
FACULTY OF MANAGEMENT University of Lodz
FEP SCHOOL OF ECONOMICS AND MANAGEMENT UNIVERSITY OF PORTO
celtech software innovations ltd
European E-Learning Institute
ACEEU ACCREDITATION COUNCIL FOR ENTREPRENEURIAL & ENGAGED UNIVERSITIES
upf. BARCELONA SCHOOL OF MANAGEMENT
Co-funded by the European Union
CLICK TO VIEW

A primary challenge André encountered was building trust in AI systems among stakeholders. Decision-makers often hesitate to fully trust AI solutions due to the perceived loss of control over decision-making. André emphasised the importance of involving stakeholders from the outset, working closely with them to validate the model and ensure it aligns with their practical domain knowledge. By doing so, stakeholders can see how the AI model complements their expertise, which helps to build trust.

In this project, one of the initial challenges involved testing assumptions about what influences discount acceptance rates. For instance, business stakeholders believed that offering more information about upcoming promotions would increase discount acceptance across all product categories. However, the AI model found that this was only true for certain products, where for others, such as meat, customers are more cautious about quality. This led to an important bias discovery, where preconceived human ideas were challenged by AI-driven insights, resulting in more accurate predictions.

Another significant challenge André highlighted was related to data quality. During the dynamic pricing project, the AI model initially produced an unexpected correlation, suggesting that higher stock levels led to greater customer acceptance of discounts. Upon further investigation, the team discovered a data tracking error within the retailer’s stock management system. This issue, which would have been overlooked without AI, was corrected, allowing for more accurate pricing strategies. The collaboration between the AI team and the business stakeholders helped refine the data and improve the model.

CONCLUSION

For André and LTP Labs, the key to successful AI adoption lies in explainability and collaboration. By involving business stakeholders early in the development process and tailoring communication to their level of analytical maturity, LTP Labs ensures that AI models are not only trusted but also optimised for practical use. For example, sensitivity analysis and partial dependence plots are used for stakeholders with limited AI proficiency, while more advanced tools like SHAP and LIME are introduced for those with deeper analytical knowledge.

The dynamic pricing project resulted in a more sustainable pricing strategy for the retailer, reducing food waste while maximising profit. AI not only improved the retailer's promotional strategies but also uncovered deeper systemic issues within the stock management process, which would have gone unnoticed without the model's insights.

Looking ahead, André believes that explainable AI will play an increasingly crucial role in both academic and business environments. As large language models (LLMs) like ChatGPT become more widespread, they will create new opportunities for making AI systems more accessible to users with varying degrees of technical expertise. In fact, André envisions a future where multi-agent AI systems interact with each other to solve complex problems more effectively, opening new doors for innovation.

> *"Preconceived human ideas were challenged by AI-driven insights, resulting in more accurate predictions."*

DIVE DEEPER

SHAP & LIME – Post-hoc explanation tools used by LTP Labs to help stakeholders understand AI model behaviours and build trust in the system.

Markdown analytics to reduce food waste at grocery stores - A case study on improving pricing strategies.

Sonae MC Adopts Analytical Model To Fight Food Waste - Press release.

Genetic Programming – Used in LTP Labs' TRUST-AI project to generate interpretable mathematical expressions that help stakeholders understand AI

Co-funded by the European Union

This project has been funded with support from the European Commission. The author is solely responsible for this publication (communication) and the Commission accepts no responsibility for any use may be made of the information contained therein. In compliance of the new GDPR framework, please note that the Partnership will only process your personal data in the sole interest and purpose of the project and without any prejudice to your rights.

FAIRNESS

Ensuring Ethical AI Development in Engineering Education and Industry

SETTING

Sašo Karakatič, PhD, is an associate professor at the University of Maribor in Slovenia, where he teaches engineering students. As a key member of the Intelligence Systems Laboratory, his research and teaching focus on AI fairness and transparency, especially in technical and advanced AI topics.

His courses are mainly designed for Master's and PhD students, ensuring that they gain both a technical understanding of AI systems and an awareness of the ethical implications these systems can create.

Karakatič's work extends beyond teaching; he is actively involved in industry projects at the national level, collaborating with Slovenian companies to integrate AI into various operations. For example, he has worked on AI projects to optimise meal distribution for companies and detect leaks in city water pipelines.

His experience also includes European-funded international projects, where AI plays a crucial role in solving broader societal challenges.

aileaders-project.eu
leaders
FACULTY OF MANAGEMENT University of Lodz
FEP
SCHOOL OF ECONOMICS AND MANAGEMENT UNIVERSITY OF PORTO
software innova
European E-Learning Institute
ACEEU
ACCREDITATION COUNCIL FOR ENTREPRENEURIAL & ENGAGED UNIVERSITIES
upf.
BARCE SCHOOL MANAGEMEN
Co-funded by the European Union
CLICK TO VIEW

> *"The most efficient way is not always the most appropriate."*

CHALLENGE

One of Karakatič's key challenges lies in addressing the unfairness that can arise in AI systems. "Unfairness stems from the data," he explains, using Amazon's infamous AI hiring tool as a prime example. The system, which was designed to scan CVs and select candidates, produced biased, sexist results. "It was sexist because the humans making decisions before were sexist," Karakatič notes. The tool simply learned from the historical data it was given, perpetuating human biases that had existed in the hiring process.

In many cases, AI models reflect the biases inherent in the data they are trained on, making fairness a complex challenge to solve. Karakatič stresses that there are opportunities to correct or "force" algorithms to be more fair. His team, for instance, is researching methods to make machine learning algorithms fairer than human decision-making processes. "AI may trigger a check on the outputs/data where checks were not happening before," he says, noting that the mere introduction of AI can sometimes prompt greater scrutiny than traditional human decisions ever received.

A second challenge is transparency in AI systems. While tools such as SHAP and LIME can help explain the decision-making processes of AI models, Karakatič points out that these tools are still largely inaccessible to non-experts. "These are programming libraries – they're useful for us researchers, but not for the average person," he explains. This creates a barrier to understanding, as the explainability metrics provided by these tools are often too technical for general users. "We're still not there in this explainability field of AI," Karakatič admits, although he remains optimistic that future developments will make these tools more widely usable.

aileaders-project.eu
leaders
FACULTY OF MANAGEMENT University of Lodz
FEP SCHOOL OF ECONOMICS AND MANAGEMENT UNIVERSITY OF PORTO
feltech software innovations ltd
European E-Learning Institute
ACEEU ACCREDITATION COUNCIL FOR ENTREPRENEURIAL & ENGAGED UNIVERSITIES
upf. BARCELONA SCHOOL OF MANAGEMENT
Co-funded by the European Union
CLICK TO VIEW

Karakatič also emphasises the importance of diverse perspectives in AI development. He warns against allowing engineers to work in isolation, as their focus on efficiency and problem-solving can overshadow broader ethical considerations. "Don't let us engineers work on our own," he advises. Involving professionals from fields like ethics, law, and social sciences is critical to ensuring that AI systems are not only efficient but also fair and socially responsible.

"Don't let us engineers work on our own."

“Unfairness stems from the data.”

CONCLUSION

Karakatič believes that the future of ethical AI lies in interdisciplinary collaboration. By involving experts from different fields, engineers can be made more aware of the ethical, legal, and social dimensions of AI. "The most efficient way is not always the most appropriate," he notes, underscoring the importance of balancing technical performance with ethical responsibility.

To educate the next generation of AI developers, Karakatič uses practical tools to demonstrate the potential risks of biased AI models. One of his go-to tools for non-experts is Teachable Machines by Google, which allows students to see how AI models can be misled by incomplete or biased data. By using simple, interactive tools like this, Karakatič helps his students grasp the importance of fairness and transparency in AI from an early stage in their education.

In his technical courses, Karakatič also introduces advanced tools such as FairLearn and DALex to engineering students, giving them hands-on experience with fairness and transparency methods in AI. These tools are used to highlight the complexity of ethical AI, ensuring that future AI developers understand both the technical and ethical challenges they will face.

While progress is being made, Karakatič remains realistic: “Some of my students will take these ethical considerations seriously when they enter the industry, but not all of them will.” Despite this, he is encouraged by the increasing attention to fairness and transparency within the AI community and hopes that regulations, like the EU AI Act, will continue to push companies and developers to prioritise responsible AI.

DIVE DEEPER

Teachable Machines by Google – A tool Karakatič uses to demonstrate how AI models can make mistakes when trained on limited data.

DALex – A tool for model explainability, helping users understand how decisions are made by AI models. DALex

FairLearn – A Python library designed to improve fairness in AI systems. FairLearn

IBM AI Fairness 360 – A toolkit designed to detect and mitigate bias in AI models. AI Fairness 360

SHAP – A library used to explain individual predictions of AI models. SHAP

LIME – A tool used to explain AI predictions, making models more interpretable. LIME

What-If Tool – A tool by Google for investigating model performance and fairness. What-If Tool

Co-funded by the European Union

This project has been funded with support from the European Commission. The author is solely responsible for this publication (communication) and the Commission accepts no responsibility for any use may be made of the information contained therein. In compliance of the new GDPR framework, please note that the Partnership will only process your personal data in the sole interest and purpose of the project and without any prejudice to your rights.

TRANSPARENCY & EXPLAINABILITY

Addressing the Ethical Challenges of AI Use in Academic Settings at Pontifícia Universidade Católica do Rio Grande do Sul.

SETTING

A central part of Rafael's work includes managing the **Apple Developer Academy**, a programme designed to help students and startups transform ideas into products. Chanin's involvement with AI primarily revolves around its educational applications and support for entrepreneurial ventures.

While AI use is still in the early stages at the university, Chanin notes that discussions on AI's role in education have gained traction, particularly after the introduction of tools like **ChatGPT.**

The university also houses a dedicated hub for AI-related startups, where companies work to integrate AI into their solutions. Although the university is making strides in teaching AI and supporting AI-related ventures, it is still grappling with establishing clear ethical guidelines and usage policies.

> Rafael Chanin, PhD, is an adjunct professor at the Pontifícia Universidade Católica do Rio Grande do Sul, Brazil. He teaches computer science and oversees the startup ecosystem within the university's Scientific and Technology Park.
>
> *"If you depend too much on tools to do your job, you're going to be in trouble."*

aileaders-project.eu
leaders
FACULTY OF MANAGEMENT University of Lodz
FEP SCHOOL OF ECONOMICS AND MANAGEMENT UNIVERSITY OF PORTO
software innova
European E-Learning Institute
ACEEU ACCREDITATION COUNCIL FOR ENTREPRENEURIAL & ENGAGED UNIVERSITIES
upf. BARCE SCHOOL MANAGEMEN
Co-funded by the European Union
CLICK TO VIEW

CHALLENGE

One of the most pressing challenges Chanin faces is how students are using AI tools like ChatGPT to complete assignments. He encourages the use of AI to enhance learning, but stresses that students must be critical of the outputs. Many students “just throw the question into the tool and take the output as truth,” often without analysing it. This has led to numerous instances where students submit incorrect work because they failed to provide sufficient context for the AI to generate accurate answers.

Additionally, Chanin highlights a growing issue with the traceability of AI-generated content. Since AI tools create unique outputs for each query, it becomes difficult to determine whether a student has used AI to complete their work. He describes this as an "ethical dilemma rather than a plagiarism dilemma," as AI-generated content is original but lacks human intellectual input. Chanin believes that if students use AI tools, they should either paraphrase the content or cite their use of the technology.

aileaders-project.eu
leaders
FACULTY OF MANAGEMENT
University of Lodz
FEP
SCHOOL OF ECONOMICS AND MANAGEMENT UNIVERSITY OF PORTO
feltech
software innovations ltd
European E-Learning Institute
ACEEU
ACCREDITATION COUNCIL FOR ENTREPRENEURIAL & ENGAGED UNIVERSITIES
upf. BARCELONA SCHOOL OF MANAGEMENT
Co-funded by the European Union
CLICK TO VIEW

Another challenge is resistance from other faculty members regarding AI’s role in education. While Chanin advocates for responsible use of AI, some colleagues believe that AI tools should be banned altogether. He counters this by arguing that AI tools cannot be realistically banned, given their widespread use. Instead, the focus should be on teaching students how to use AI responsibly without bypassing the critical thinking process.

CONCLUSION

To address these challenges, Chanin suggests a two-step approach for students using AI tools.

First, provide sufficient context. When using AI tools, students should ensure they give detailed context in their queries, as many AI tools struggle with answering more complex questions without it. **“There are a lot of implicit material content or variables you take for granted, but all of these 'obvious' things have to be included in the query,”** he explains

Secondly, double check AI outputs. Students must critically evaluate the output generated by AI, especially when it involves tasks like coding or calculations. Chanin emphasizes that **“AI tools don’t necessarily test the output,”** so students should take responsibility for verifying the accuracy of the information provided.

Chanin also believes that AI developers have an ethical obligation to make their decision-making processes more transparent. He acknowledges the complexity of this issue, particularly with private companies that are not required to share their proprietary code. However, he suggests that developers should at least explain how their tools generate results, even if they cannot fully open their systems for public scrutiny. This transparency would help users better understand potential biases and limitations.

Furthermore, Chanin is optimistic about AI’s potential for future integration across various platforms. He is particularly impressed by how tools like **Microsoft Teams** and **Google Meet** are already using AI to generate insights from meetings, such as next steps or behavioural analyses. **“These tools might soon be able to predict what we want to do next,”** he notes, as AI becomes increasingly integrated into daily life. In his message to students, Chanin stresses the importance of “learning how to learn.”

While shortcuts provided by AI tools can be useful, he warns that over-reliance on them can lead to long-term challenges. He urges students to use AI as a learning aid rather than a crutch, stating, **“If you depend too much on tools to do your job, you’re going to be in trouble”**.

PUCRS
Pontifícia Universidade Católica do Rio Grande do Sul

DIVE DEEPER

Apple Developer Academy - A programme within the university’s tech park aimed at helping students and startups.

The NAVI Hub - A hub within the university's tech park focused on AI-related startups

AI for Microsoft Teams

AI for Google Meet

Co-funded by the European Union

This project has been funded with support from the European Commission. The author is solely responsible for this publication (communication) and the Commission accepts no responsibility for any use may be made of the information contained therein. In compliance of the new GDPR framework, please note that the Partnership will only process your personal data in the sole interest and purpose of the project and without any prejudice to your rights.

ETHICS

Enhancing Surgical Decisions: Balancing Innovation and Ethics in AI-Driven Healthcare

SETTING

Ciaran McCourt, originally from Ireland and now based in the United States, has over 20 years of experience in bringing European startups, primarily in the MedTech and digital health sectors, into the US market. His latest role is as CEO of Move Up, a Belgian-based health technology company. The company, founded by an orthopaedic surgeon and a team of engineers, specialises in collecting perioperative patient data—information gathered before and after surgery. Move Up aims to provide surgeons with a clearer picture of a patient's health status, offering clinical insights not previously available. This data is gathered from a range of sources, including wearable sensors and digital applications, to assess factors like pain levels, physical activity, and overall well-being. Additionally, Move Up has incorporated AI-based predictive models to evaluate whether a patient is a suitable candidate for surgery.

> *"Ultimately, we've stayed on the side of it being a tool for the surgeons to make better decisions... It's kind of an education tool. We don't ever go as far as to say this tool makes the decision."*

CHALLENGE

aileaders-project.eu
leaders
FACULTY OF MANAGEMENT
University of Lodz
FEP
SCHOOL OF ECONOMICS AND MANAGEMENT
UNIVERSITY OF PORTO
celtech
software innovations ltd
European E-Learning Institute
ACEEU
ACCREDITATION COUNCIL FOR ENTREPRENEURIAL & ENGAGED UNIVERSITIES
upf.
BARCELONA SCHOOL OF MANAGEMENT
Co-funded by the European Union
CLICK TO VIEW

The implementation of AI in medical decision-making raises significant ethical and practical challenges. One key challenge identified by McCourt revolves around predictive modelling—using AI to forecast patient outcomes. This involves building a patient "phenotype" by collecting subjective and objective data. AI is then used to predict how patients with similar profiles have responded to procedures like knee or hip implants. This predictive approach helps surgeons determine whether surgery is appropriate and the likelihood of a successful outcome. This has the potential to improve patient care while discussing new ethical questions, particularly around decision autonomy. Surgeons may one day face dilemmas such as whether AI should directly inform surgical decisions or merely serve as a supportive tool.

Another major challenge is the acceptance of these AI tools by medical professionals, particularly surgeons who are often sceptical about software solutions. The technology is promising, the integration of AI into daily workflows remains a barrier. Many surgeons are accustomed to established processes and resist additional layers of software that may disrupt existing practices.

This issue is exacerbated by the reluctance to integrate AI tools with existing Electronic Health Records (EHR) systems, which are often dominated by major players like Epic, known for their complex interfaces.

The financial aspect presents a broader ethical consideration, particularly in the US healthcare system, where insurance policies heavily influence healthcare decisions. AI tools being developed across the industry, have potential implications for insurance providers, as they can predict the necessity of surgery and postoperative care. This may lead to suggestions that AI could inadvertently influence medical decisions, potentially prioritising efficiency or cost-saving over patient needs.

CONCLUSION

To address these challenges, Move Up has developed AI-driven tools that are intended to support, not replace, the decision-making process of surgeons. By providing comprehensive patient profiles and predictive insights, the AI system aims to optimise consultation time, identify suitable candidates for surgery, and reduce unnecessary procedures. This creates a dual benefit—enhancing patient care while increasing efficiency in hospital settings. Move Up's approach has been well-received by surgeons open to technological innovation, particularly those in private practices who manage their own patient lists.

Looking forward, the company plans to expand its AI applications to the pharmaceutical sector, focusing on real-world data collection to assess the effectiveness of medications and adherence to treatment protocols. This could provide pharmaceutical companies with invaluable insights into patient outcomes and overall quality of life. Although Move Up's AI tools are still in the early stages of adoption, they show significant promise in reducing costs, improving care efficiency, and offering predictive insights that enhance clinical decision-making.

moveUP

"AI will not get rid of doctors per se, but will get rid of doctors who do not use AI."

DIVE DEEPER

moveUP - Digital platforms for smarter, more efficient healthcare delivery

AI Conferences: often focus on AI ethics, digital health, and responsible AI. These gatherings offer insights into the latest trends in AI and healthcare.

- **AI for Good - Global Summit 2024**
- **Google Cloud Next - '24 Opening Keynote**
- **Reuters Momentum Summit - explores the future of AI innovation**

Epic Systems: Electronic Health Records (EHR) provider crucial in understanding the complexities of data management in healthcare.

Co-funded by the European Union

This project has been funded with support from the European Commission. The author is solely responsible for this publication (communication) and the Commission accepts no responsibility for any use may be made of the information contained therein. In compliance of the new GDPR framework, please note that the Partnership will only process your personal data in the sole interest and purpose of the project and without any prejudice to your rights.

FAIRNESS

Balancing Automation with Human Oversight in AI-Driven Marketing

SETTING

Jesper Valentin Holm is the Chief Executive Officer (CEO) of Cobiro, a company specialising in AI-driven solutions within the marketing technology (MarTech) sector. Cobiro's AI-powered tools are designed to improve decision-making and automate digital advertising campaigns for small to medium-sized enterprises (SMEs).

Cobiro operates across various European markets, using AI to simplify and optimise digital marketing processes, making these advanced technologies more accessible to businesses that lack deep technical expertise.

With over 15 years of experience in technology leadership, Holm's expertise lies in leveraging AI to drive efficiency and innovation in digital marketing. Cobiro focuses on using AI to automate decision-making, optimise ad campaigns, and reduce costs for SMEs, enabling them to compete more effectively in the online advertising space.

> *"AI should solve real business needs, not just be adopted for the sake of technology."*

aileaders-project.eu
leaders
FACULTY OF MANAGEMENT University of Lodz
FEP
SCHOOL OF ECONOMICS AND MANAGEMENT UNIVERSITY OF PORTO
software innova
European E-Learning Institute
ACEEU
ACCREDITATION COUNCIL FOR ENTREPRENEURIAL & ENGAGED UNIVERSITIES
upf. BARCE SCHOOL MANAGEMEN
Co-funded by the European Union

CLICK TO VIEW

CHALLENGE

One of the main challenges Holm faced was ensuring fairness and transparency in Cobiro's AI systems, especially in digital advertising where biased outcomes could unfairly target or exclude certain user groups. Cobiro proactively addresses this by regularly monitoring and testing their AI algorithms to identify any potential biases in ad placements. "We focus on continuous monitoring to detect any patterns or anomalies across different demographics and user segments," Holm explains. Cobiro has built processes to track how AI decisions impact diverse groups and make adjustments as needed to maintain balanced and fair outcomes.

Holm also highlighted the importance of transparency and explainability in AI. He believes that users should not only be able to use AI but also understand how it works. “You don’t need to be a data scientist to understand how an algorithm works,” he explains, “but you do need to understand the core concepts to make informed decisions.” At Cobiro, AI decision-making processes are continuously reviewed and made more transparent, with feedback loops established to gather user input. Holm believes that providing clear explanations of AI’s functionality is essential, particularly in the context of businesses that rely on these systems for marketing decisions.

aileaders-project.eu

AI leaders

FACULTY OF MANAGEMENT University of Lodz

FEP SCHOOL OF ECONOMICS AND MANAGEMENT UNIVERSITY OF PORTO

feltech software innovations ltd

European E-Learning Institute

ACEEU ACCREDITATION COUNCIL FOR ENTREPRENEURIAL & ENGAGED UNIVERSITIES

upf. BARCELONA SCHOOL OF MANAGEMENT

Co-funded by the European Union

CLICK TO VIEW

Another challenge is the balance between automation and user control. While Cobiro's AI helps automate repetitive tasks, such as optimising bids for ad campaigns or identifying high-performing keywords, Holm recognises that human oversight remains critical. "Our AI can give suggestions, but it's the users who ultimately decide what direction to take," Holm notes. He acknowledges that AI lacks the contextual understanding that humans provide, meaning its outputs must be interpreted and adjusted by human users.

Lastly, ethical concerns around data privacy and security are ever-present. Although Cobiro primarily works with transactional rather than personal data, the company complies with GDPR and similar regulations, ensuring that any data it collects is securely processed and users are informed about how their data is used.

CONCLUSION

Holm sees AI as a tool that augments human decision-making rather than replacing it. He advises businesses to start small when implementing AI and be agile in iterating their systems. "Focus on one problem at a time and iterate quickly. AI should solve real business needs, not just be adopted for the sake of technology," Holm advises. He also emphasises the importance of engaging diverse perspectives, both in terms of the development team and user feedback, to ensure that AI systems are both fair and transparent.

Cobiro's AI has proven to be a significant advantage for SMEs, automating complex tasks like ad placement optimisation and allowing these businesses to compete more effectively in the digital marketing space. By improving efficiency, Cobiro has helped make marketing more accessible and affordable for businesses that previously lacked the resources to engage in sophisticated advertising strategies.

Looking to the future, Holm is excited about advancements in Natural Language Processing (NLP), which will allow AI systems to interact with users more naturally and interpret unstructured data more effectively. He also sees predictive AI becoming a key player in marketing technology, enabling businesses to anticipate future outcomes based on data analysis and further refine their marketing strategies.

> *“You don’t need to be a data scientist to understand how an algorithm works, but you do need to understand the core concepts to make informed decisions.”*

DIVE DEEPER

Jupyter Notebooks and Google Colab – Tools recommended for teaching AI concepts, enabling students to experiment with AI models through hands-on learning.

Uniscale – A startup developing tools to translate business requirements into technical specifications, helping with rapid prototyping and AI implementation.

Co-funded by the European Union

This project has been funded with support from the European Commission. The author is solely responsible for this publication (communication) and the Commission accepts no responsibility for any use may be made of the information contained therein. In compliance of the new GDPR framework, please note that the Partnership will only process your personal data in the sole interest and purpose of the project and without any prejudice to your rights.

FAIRNESS

Optimising Product Data and Navigating in E-commerce with AI

SETTING

Morten Poulsen is the Founder and CEO of Plytix, a company specialising in product information management (PIM) software, aimed at helping businesses streamline their multichannel retail operations. Plytix provides tools for automating content creation, product optimisation, and data management across various e-commerce platforms. With years of experience in the AI-driven optimisation of product information, Morten's work focuses on leveraging AI to enhance content creation, data enrichment, and workflow efficiencies in the e-commerce and marketing sectors.

Operating at an international level, Plytix serves businesses globally, particularly in the European market. The company's AI tools are designed to help small to medium-sized enterprises (SMEs) manage and optimise their product data for multiple online sales channels, thereby improving their ability to compete in the digital marketplace.

> *"We make sure that our users understand how the AI works and why it makes certain recommendations."*

aileaders-project.eu
leaders
FACULTY OF MANAGEMENT University of Lodz
FEP SCHOOL OF ECONOMICS AND MANAGEMENT UNIVERSITY OF PORTO
software innova
European E-Learning Institute
ACEEU ACCREDITATION COUNCIL FOR ENTREPRENEURIAL & ENGAGED UNIVERSITIES
upf. BARCE SCHOOL MANAGEMEN
Co-funded by the European Union

Co-funded by the European Union. Views and opinions expressed are however those of the author or authors only and do not necessarily reflect those of the European Union or the Foundation for the Development of the Education System. Neither the European Union nor

CLICK TO VIEW

CHALLENGE

One of the primary challenges that Poulsen and Plytix tackle is optimising product information management for businesses operating across multiple platforms. The task of creating, maintaining, and optimising product listings for different sales channels—such as Amazon, Google Shopping, and other e-commerce marketplaces—can be daunting, particularly for SMEs with limited resources. "If you're managing thousands of product listings across 20 different marketplaces, it's impossible to manually optimise each one," Poulsen explains. AI plays a crucial role in automating this process, generating platform-specific product descriptions and titles, and ensuring that businesses meet the unique requirements of each marketplace.

Plytix uses AI to streamline these tasks, allowing businesses to input product data into the system, which then optimises and generates descriptions tailored for each platform. Poulsen highlights the importance of using clean data for AI to generate accurate and high-quality product information. "If you feed bad data into an AI, you'll get bad results out. That's why we ensure the data going in is clean," he notes. Plytix's AI-driven tools help businesses manage these processes at scale, improving both productivity and the quality of product listings.

A key challenge that Morten Poulsen emphasises is fairness and bias in AI systems. To address this, Plytix ensures that its AI models are designed with diverse data inputs and undergo continuous testing to avoid biased outcomes. "We track the AI's outputs across different datasets to ensure there are no skewed patterns or unfair recommendations," says Poulsen.

aileaders-project.eu
leaders
FACULTY OF MANAGEMENT University of Lodz
FEP
SCHOOL OF ECONOMICS AND MANAGEMENT UNIVERSITY OF PORTO
feltech software innovations ltd
European E-Learning Institute
ACEEU ACCREDITATION COUNCIL FOR ENTREPRENEURIAL & ENGAGED UNIVERSITIES
upf. BARCELONA SCHOOL OF MANAGEMENT
Co-funded by the European Union
CLICK TO VIEW

For example, Plytix identified skewed recommendations based on overly narrow datasets. By diversifying data inputs, they were able to correct the biases and produce more balanced recommendations.

In terms of transparency and accountability, Poulsen stresses the importance of clear communication and explainability in AI systems. Plytix integrates feedback mechanisms into its platform, enabling users to provide real-time feedback on AI performance. Based on this feedback, the company adjusts algorithms to better align with user needs and ensure that the AI remains transparent and reliable. "We make sure that our users understand how the AI works and why it makes certain recommendations," Poulsen explains, noting that providing transparency is essential for building user trust.

Another significant challenge is data privacy and regulatory compliance, particularly under the General Data Protection Regulation (GDPR) in the European Union. Plytix handles sensitive data responsibly by using encryption, ensuring secure data storage, and complying with privacy laws. "We prioritise user privacy by building GDPR compliance into the very core of our AI systems," Poulsen affirms, adding that data minimisation and informed consent are critical components of their approach.

CONCLUSION

AI has become a key driver of efficiency and innovation at Plytix, particularly in automating the manual tasks associated with managing and optimising product data across e-commerce platforms. Poulsen's philosophy is to focus on specific problems when implementing AI, ensuring that the solutions provide clear value and measurable results. He advises businesses to start small and iterate quickly, testing AI tools in real-world scenarios before scaling them up.

Poulsen is optimistic about the future of AI, especially in areas like natural language processing (NLP) and predictive analytics, which he believes will further revolutionise product data management by allowing businesses to personalise and enhance their offerings more intuitively. "As AI evolves, we expect it to transform how businesses handle product data, making it more intuitive and efficient for everyone," Poulsen explains.

Through the integration of diverse data inputs, bias detection mechanisms, and transparent communication channels, Plytix ensures that its AI systems deliver fair, reliable, and efficient outcomes for businesses. As AI continues to advance, Poulsen remains committed to maintaining ethical standards and user trust, emphasising that AI should always be developed with a focus on fairness, transparency, and privacy.

> *"We prioritise user privacy by building GDPR compliance into the very core of our AI systems."*

DIVE DEEPER

Natural Language Processing (NLP) – Plytix is exploring NLP technologies to further personalise and enhance product data management, allowing businesses to interact more naturally with their data systems.

Product Information Management (PIM) Software – The core tool developed by Plytix to manage, optimise, and distribute product data across multiple e-commerce platforms.

Plytix - specialising in product information management (PIM) software

BY SA

Co-funded by the European Union

This project has been funded with support from the European Commission. The author is solely responsible for this publication (communication) and the Commission accepts no responsibility for any use may be made of the information contained therein. In compliance of the new GDPR framework, please note that the Partnership will only process your personal data in the sole interest and purpose of the project and without any prejudice to your rights.

23

ETHICS

Balancing Human Creativity and AI Efficiency in Business Operations

SETTING

Roan Van Der Sluis has been with AI Heroes for two and a half years, serving in various leadership roles, where he divides his time between company management (30%) and direct project and team management (70%). AI Heroes specialises in integrating and implementing AI solutions to streamline business operations, optimise processes, and drive innovation across different sectors. Roan's work focuses on applying generative AI to automate routine tasks, allowing employees to concentrate on more creative and meaningful work.

AI Heroes operates primarily in Europe and engages in projects with wide-reaching impacts, such as a pilot with a Dutch institute. A key area of focus is the company's Base project, which uses AI to automate customer support services, enabling businesses to scale efficiently while maintaining quality.

> *"AI handles the mundane tasks, freeing up time for people to do the parts of their job they enjoy the most."*

aileaders-project.eu

leaders
FACULTY OF MANAGEMENT University of Lodz
FEP
SCHOOL OF ECONOMICS AND MANAGEMENT UNIVERSITY OF PORTO
feltrero software innovation
European E-Learning Institute
ACEEU ACCREDITATION COUNCIL FOR ENTREPRENEURIAL & ENGAGED UNIVERSITIES
upf. BARCELONA SCHOOL OF MANAGEMENT
Co-funded by the European Union

CLICK TO VIEW

CHALLENGE

Roan stresses the importance of using both an AI-first and AI-second approach, depending on the complexity and type of task. In an AI-first approach, AI is the primary solution for handling tasks like routine content generation or initial problem-solving, particularly when tackling unfamiliar areas. “If you're working on something outside your expertise, AI can help speed things up and streamline the process,” Roan explains. However, for more nuanced or creative work, the AI-second approach is more effective. Here, human expertise drives the task, with AI providing supplementary support. Roan gives the example of using AI to generate initial marketing copy, which was later refined by human input to ensure quality and depth.

One of the critical ethical concerns Roan faces is the potential for bias in AI systems. AI Heroes actively addresses bias in their development processes, including through a recent project aimed at reducing biases in interview notes. Roan acknowledges the difficulty of completely eliminating bias but emphasises that AI Heroes incorporates bias detection training for developers to help mitigate these issues. They regularly monitor AI outputs and make real-world adjustments to ensure that bias does not affect outcomes, while maintaining control over the training and bias detection processes.

In addition to bias, Roan points to the fear of job displacement as a recurring issue when implementing AI-driven solutions. However, he asserts that AI has not led to job losses within AI Heroes. Instead, AI allows employees to move away from repetitive tasks, focusing on more creative and impactful work. "AI handles the mundane tasks, freeing up time for people to do the parts of their job they enjoy the most," Roan explains. He also recognises that generative AI will likely be as transformative as the industrial revolution, fundamentally altering how businesses operate.

aileaders-project.eu
leaders
FACULTY OF MANAGEMENT University of Lodz
FEP
SCHOOL OF ECONOMICS AND MANAGEMENT UNIVERSITY OF PORTO
feltech software innovations ltd
European E-Learning Institute
ACEEU
ACCREDITATION COUNCIL FOR ENTREPRENEURIAL & ENGAGED UNIVERSITIES
upf. BARCELONA SCHOOL OF MANAGEMENT
Co-funded by the European Union
CLICK TO VIEW

Another challenge revolves around data privacy in AI projects, particularly in sensitive use cases like elderly care. AI Heroes developed a fall detection system that uses computer vision technology to help elderly people live independently at home. To address privacy concerns, they implemented edge devices that process data locally, ensuring no personal information leaves the home unless an emergency occurs. Roan stresses that privacy protection is a core part of AI Heroes' development strategy, especially in projects involving vulnerable populations.

> *“This concentration of power is one of the biggest ethical issues our society faces today.”*

CONCLUSION

Roan believes that AI is not about replacing humans but rather about enhancing their capabilities. By automating repetitive tasks, AI Heroes has allowed employees to focus on creative and strategic work, which adds more value to the business. This aligns with Roan's broader vision for AI, which he believes will democratise tasks, enabling individuals to perform duties beyond their expertise. "AI allows me to do tasks that I'm not naturally skilled at, such as generating graphics or analysing data, which I can then use to improve my work," he explains.

In terms of societal impact, AI Heroes' solutions, such as their customer support automation tools and fall detection system for elderly care, have helped businesses improve operational efficiency and user experience. While these advancements bring clear benefits, Roan acknowledges the ethical challenges posed by concentration of AI power in the hands of a few large companies. "This concentration of power is one of the biggest ethical issues our society faces today," he notes, pointing to the need for fair and transparent AI systems.

AI Heroes also takes proactive steps to ensure ethical oversight throughout their projects. Before development begins, discussions are held to assess the potential impact of AI solutions on privacy, fairness, and societal values. This continuous ethical reflection ensures that AI solutions are designed responsibly, with both risks and benefits in mind.

> *"AI allows me to do tasks that I'm not naturally skilled at, such as generating graphics or analysing data, which I can then use to improve my work."*

DIVE DEEPER

AI First and AI Second Strategies – AI Heroes uses these strategies to balance automation with human input, ensuring efficiency without sacrificing quality.

Base by AI Heroes – A generative AI project that automates customer support services, improving scalability for businesses.

Edge Devices for Privacy – Used in fall detection systems to process data locally, preserving privacy while providing critical services.

Co-funded by the European Union

This project has been funded with support from the European Commission. The author is solely responsible for this publication (communication) and the Commission accepts no responsibility for any use may be made of the information contained therein. In compliance of the new GDPR framework, please note that the Partnership will only process your personal data in the sole interest and purpose of the project and without any prejudice to your rights.

SOCIETAL IMPACT

Leveraging AI for Reliable and Secure Emissions Reporting

SETTING

Chris Walton is the co-founder of Tool Zero, a startup focused on helping small and medium-sized businesses track, report, and take action on their greenhouse gas emissions. Tool Zero uses AI technology to automate emissions reporting, integrating various AI tools to simplify the complex process of extracting, calculating, and presenting emission data.

Tool Zero leverages AI in several key areas:

- Semantic similarity search, allowing users to search not just by keywords but by the meaning of the terms. This is made possible by large language models that excel at understanding the semantics of text.

- Data extraction from documents, combining optical character recognition (OCR) with AI to extract meaningful information such as invoice numbers from scanned PDFs and images.

- Generative outputs, where once emissions data is calculated, AI helps generate reports tailored to specific reporting frameworks, such as regulatory disclosures or the Carbon Disclosure Project.

Walton's experience in using AI focuses on its ability to streamline emissions tracking for businesses, providing actionable insights from large volumes of data without requiring extensive manual input.

> *“We ensure that AI models do not train on customer data, maintaining trust and data security.”*

aileaders-project.eu

CLICK TO VIEW

leaders
FACULTY OF MANAGEMENT University of Lodz
FEP SCHOOL OF ECONOMICS AND MANAGEMENT UNIVERSITY OF PORTO
software innovati
European E-Learning Institute
ACEEU ACCREDITATION COUNCIL FOR ENTREPRENEURIAL & ENGAGED UNIVERSITIES
upf. BARCE SCHOOL MANAGEM
Co-funded by the European Union

aileaders-project.eu
leaders
FACULTY OF MANAGEMENT University of Lodz
FEP SCHOOL OF ECONOMICS AND MANAGEMENT UNIVERSITY OF PORTO
feltech software innovations ltd
European E-Learning Institute
ACEEU
upf. BARCELONA SCHOOL OF MANAGEMENT
CLICK TO VIEW

CHALLENGE

One of the main challenges Walton highlights is the reliability of AI. Large language models (LLMs), which power many AI solutions, are probabilistic rather than deterministic. This means they do not always return the same output given the same input, which can lead to inconsistencies. Walton describes this as a “fundamental challenge,” where the AI can sometimes produce hallucinations—incorrect or unpredictable outputs that don’t align with expectations.

To address this, Tool Zero has developed a system that bounds the output space. By limiting the range of acceptable outputs, they reduce the chances of AI providing incorrect or nonsensical results. Walton explains that this approach works particularly well for Tool Zero’s use case, where emissions calculations are confined to a known set of possible outputs. However, he acknowledges that this challenge is ongoing, and there is still work to be done to ensure AI reliability across all use cases.

Another key ethical challenge Tool Zero faces is data security and privacy. Walton points out the importance of ensuring that sensitive business information, such as purchase invoices and proprietary documents, is handled securely. A significant concern is that AI models, especially those hosted by third parties like OpenAI or Google, might inadvertently train on this data, leading to unintentional data leakage. Walton references a past case where proprietary code uploaded to an AI model was later regurgitated to an unauthorised user—a scenario that Tool Zero is keen to avoid.

Tool Zero mitigates this risk by ensuring that AI models do not train on customer data. Additionally, Walton emphasises the importance of secure cloud infrastructure and robust data handling protocols to maintain customer trust. This is particularly important when working with emissions data that could reveal sensitive information about a company’s operations and purchasing habits.

> *“AI can sometimes produce hallucinations—incorrect or unpredictable outputs that don’t align with expectations.”*

DIVE DEEPER

Semantic Similarity Search – Tool Zero uses this to allow searches based on the meaning of words rather than just keywords, powered by large language models.

Toolzero - Calculate Your Scope 1, 2, and 3 Emissions in Minutes.

Data Extraction with OCR – A combination of optical character recognition and AI helps extract meaningful data, such as invoice numbers, from documents.

Bounded Output Systems – This approach limits the AI's possible outputs to ensure reliability and prevent hallucinations in the generated results.

CONCLUSION

AI has played a crucial role in streamlining emissions reporting for Tool Zero's clients. By automating data extraction and report generation, businesses can quickly and efficiently calculate their greenhouse gas emissions without needing to engage deeply with the technical details. This allows small and medium-sized businesses to comply with increasingly stringent environmental regulations without investing significant resources into manual data management.

Walton is particularly optimistic about the future of semantic understanding in AI, which enables systems like Tool Zero's to extract meaning from documents and match data with emissions factors accurately. However, he acknowledges that ongoing work is required to improve AI reliability and data security, both critical for maintaining the trust and satisfaction of Tool Zero's clients.

Looking ahead, Walton sees AI as continuing to evolve and become more powerful, with new applications emerging across different industries. However, he also stresses the importance of developing robust ethical frameworks and regulations to ensure that AI systems remain accountable, fair, and transparent. Tool Zero is committed to working within these evolving standards, ensuring that their AI-driven solutions provide reliable and ethical services to their clients.

> *"By automating data extraction and report generation, businesses can quickly and efficiently calculate their greenhouse gas emissions."*

Co-funded by the European Union

This project has been funded with support from the European Commission. The author is solely responsible for this publication (communication) and the Commission accepts no responsibility for any use may be made of the information contained therein. In compliance of the new GDPR framework, please note that the Partnership will only process your personal data in the sole interest and purpose of the project and without any prejudice to your rights.

REGULATORY COMPLIANCE

Balancing Automation, Ethics, and Compliance in AI Tax Systems

SETTING

Artur Tim is a PhD candidate at the University of St. Gallen in Switzerland and serves as a tax compliance partner for the city of Munich. His background is unique, blending legal expertise with AI, as he works on automating complex processes in international tax law. Artur's experience spans both academia and industry, with notable contributions in tax automation for multinational companies like Flex Mobility (the owner of Flixbus and Greyhound in the United States), where he leads efforts to automate VAT and sales tax compliance across more than 40 countries.

Additionally, Artur previously worked at the BMW Research and Innovation Centre, training AI models. In both his professional roles and academic research, he bridges data science, data analytics, and tax law, exploring how AI can optimise tax compliance while adhering to stringent regulatory frameworks.

aileaders-project.eu
leaders
FACULTY OF MANAGEMENT University of Lodz
FEP SCHOOL OF ECONOMICS AND MANAGEMENT UNIVERSITY OF PORTO
European E-Learning Institute
ACEEU ACCREDITATION COUNCIL FOR ENTREPRENEURIAL & ENGAGED UNIVERSITIES
upf. BARCELONA SCHOOL OF MANAGEMENT
Co-funded by the European Union
CLICK TO VIEW

> *“Should AI be trusted to make decisions with significant financial or legal consequences, or should human supervision always be involved?”*

University of St.Gallen

CHALLENGE

One of the primary ethical challenges Artur faces in his work is the question of whether AI models should be allowed to make independent decisions in sensitive areas like tax compliance and law enforcement. He offers an example from Switzerland, where software systems automatically generate decisions for speeding tickets and tax penalties. While individuals can contest these decisions, many remain unaware of this right. The ethical dilemma, as Artur describes, is not whether these systems are legal but whether they are ethical. Should AI be trusted to make decisions with significant financial or legal consequences, or should human supervision always be involved?

In his view, it's crucial to integrate human oversight in such systems, particularly in areas like tax law, where errors could have long-term impacts on individuals or businesses. Artur stresses that while AI can automate many processes, it is vital to establish checks and balances to ensure that these decisions are fair and ethical.

Another major challenge Artur highlights is regulatory compliance, particularly with the General Data Protection Regulation (GDPR). GDPR, which governs data privacy and security across the European Union, is one of the most critical regulations AI developers must navigate.

aileaders-project.eu
leaders
FACULTY OF MANAGEMENT University of Lodz
FEP SCHOOL OF ECONOMICS AND MANAGEMENT UNIVERSITY OF PORTO
celtech software innovations ltd
European E-Learning Institute
ACEEU ACCREDITATION COUNCIL FOR ENTREPRENEURIAL & ENGAGED UNIVERSITIES
upf. BARCELONA SCHOOL OF MANAGEMENT
Co-funded by the European Union
CLICK TO VIEW

Artur explains that to build high-quality AI models, developers need access to large datasets, often with sensitive or non-anonymised data. However, GDPR restricts the use of such data, making it challenging to train effective AI models without breaching privacy rules.

This situation contrasts sharply with countries like China, India, and the United States, where data privacy regulations are more relaxed. As a result, AI researchers in those countries often have greater access to data, allowing them to push the boundaries of AI research more rapidly than their European counterparts. For Artur, this creates a costly burden in Europe, as adhering to GDPR compliance involves extensive resources, including legal expertise, to ensure that AI models are developed in accordance with privacy and copyright laws.

CONCLUSION

To overcome these challenges, Artur has developed a set of guidelines and rules to ensure tax compliance and ethical AI usage. In Munich, for instance, he helped design a tax compliance management system that outlines how AI processes should operate within legal frameworks. This system includes detailed documentation and auditing mechanisms that ensure the AI models remain compliant with regulations like GDPR and tax laws. He also emphasises the importance of external control, advocating for third-party oversight to ensure AI models are used correctly and ethically.

Despite the challenges of regulatory compliance, Artur is excited about the potential of generative AI. He is currently working on automating tax decisions in German-speaking countries using generative AI models like GPT. These models are already showing promising results, allowing for the efficient handling of complex tax decisions. For Artur, generative AI represents a transformative technology that will make advanced AI capabilities more accessible, even in traditionally conservative fields like tax law.

Artur also notes the growing importance of ethics in AI, particularly in countries with fewer regulatory frameworks. In Europe, where regulations are stringent, compliance takes priority, but in other regions, there is a greater need to focus on ensuring AI is used ethically. This creates an opportunity for European companies and researchers to lead the conversation on ethical AI while still complying with regulations.

> *“Adhering to GDPR compliance involves extensive resources, including legal expertise, to ensure that AI models are developed in accordance with privacy and copyright laws.”*

DIVE DEEPER

Jupyter Notebook & Anaconda – For AI model training, Artur uses Jupyter Notebook and Anaconda. These tools are popular among researchers and educators but can be difficult for students to install locally due to storage and complexity issues. Artur suggests the need for more accessible, online-based AI training platforms.

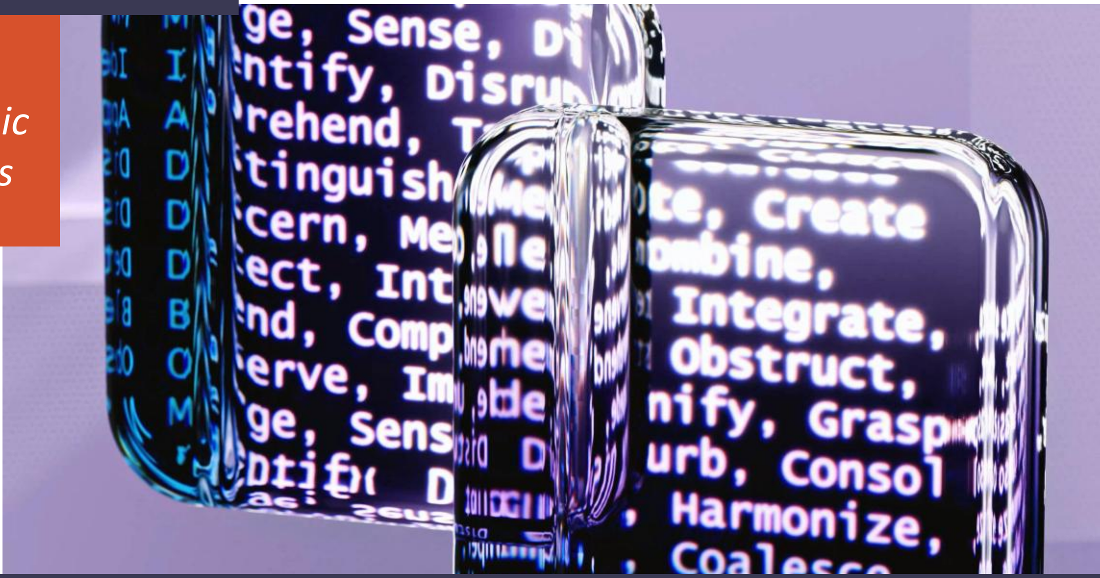
Alteryx – a tool commonly used by Big 4 firms for building AI models without requiring extensive programming knowledge. However, he pointed out that this tool can be prohibitively expensive, highlighting the need for more affordable AI training platforms.

Co-funded by the European Union

This project has been funded with support from the European Commission. The author is solely responsible for this publication (communication) and the Commission accepts no responsibility for any use may be made of the information contained therein. In compliance of the new GDPR framework, please note that the Partnership will only process your personal data in the sole interest and purpose of the project and without any prejudice to your rights.

SOCIETAL IMPACT

Bridging AI Ethics and Accountability in Academic and Professional Settings



SETTING

Damian Kędziora is an associate professor at Lahti University and adjunct faculty at the University of Warsaw in the Department of Management of Society Network. His diverse career spans more than ten years across banking, IT consulting, and software engineering. Damian has also held various roles in global organisations, starting as a programmer and transitioning into service desk management, product engineering, service management, and sales and marketing. Currently, his focus is on academia, but he continues to integrate AI models into his teaching and research.

At the university, Damian primarily uses language models like ChatGPT to support his research and teaching. While AI is not heavily utilised institution-wide, Damian's experience with low-code applications like Power Automate has proven beneficial in automating smaller tasks, enabling more efficient workflows in academic and research environments.

Damian's previous work at Norian (formerly ECIT) involved integrating AI across a wide range of services and products, incorporating components such as forecasting, anomaly detection, and recommendation systems into their portfolio. These AI components were critical in optimising various operations, particularly in IT services and consultancy, supporting a global customer base that included markets in the US, Singapore, and China.

> *"AI should be viewed as a tool to aid and not replace original student contributions."*

CLICK TO VIEW

CHALLENGE

One of the main challenges Damian identifies is accountability in AI solutions, particularly when using probabilistic software such as machine learning algorithms. Unlike deterministic software, which guarantees consistent outputs, probabilistic AI tools introduce accuracy rates—often 95% or 98%—but never 100%. Damian stresses that despite AI's high accuracy, there remains a risk of inaccuracy, potentially leading to misconduct or wrong transactions. This raises questions of responsibility and whether the organisation or a particular individual should be held accountable when AI-generated outputs fail.

Another challenge Damian discusses is the ethical use of AI, especially in academic settings. While AI tools can greatly improve efficiency in teaching and research, there is a risk of improper acknowledgement of AI contributions. Damian encourages transparency, advocating for clear declarations from students and researchers on how AI tools were used in their work. Failing to acknowledge AI use can result in ethical concerns, as individuals may falsely claim work generated by AI as their own.

aileaders-project.eu
AI leaders
FACULTY OF MANAGEMENT University of Lodz
FEP SCHOOL OF ECONOMICS AND MANAGEMENT UNIVERSITY OF PORTO
celtech software innovations ltd
European E-Learning Institute
ACEEU ACCREDITATION COUNCIL FOR ENTREPRENEURIAL & ENGAGED UNIVERSITIES
upf. BARCELONA SCHOOL OF MANAGEMENT
Co-funded by the European Union
CLICK TO VIEW

Furthermore, Damian highlights the ongoing challenge of bridging theory and practice in education. While AI can assist students in polishing their work, referencing properly, and searching literature, the creative work must ultimately come from the students themselves. He emphasises that AI should be viewed as a tool to aid and not replace original student contributions.

> *“Probabilistic AI models should always come with a clear understanding of their limitations.”*

CONCLUSION

Damian's solution to managing AI's potential for error and accountability focuses on the importance of organisational responsibility. Companies, universities, or organisations employing AI should retain full accountability for the outputs of their systems, ensuring that liability does not rest solely on the algorithm. Damian believes that probabilistic AI models should always come with a clear understanding of their limitations, and organisations must establish liability clauses to address inaccuracies when they arise.

On the topic of ethical AI use, Damian is a strong advocate for honesty and transparency in both educational and professional settings. He encourages universities and organisations to implement policies requiring clear declarations of AI use. At his university, for example, all student theses must now include an AI user declaration, providing transparency in how AI was applied. Damian views this as a critical step toward ensuring ethical AI integration in academic work.

Regarding automation, Damian remains a strong proponent of using AI and low-code tools to reduce waste, time, and costs. He has embraced the use of language models in teaching, encouraging students to practice effective prompting techniques to maximise the value AI offers. Despite some minor resistance, the overwhelming majority of feedback he receives is positive, with students recognising the benefits of AI in streamlining their work.

> *"Organisations employing AI should retain full accountability for the outputs of their systems, ensuring that liability does not rest solely on the algorithm."*

DIVE DEEPER

ChatGPT – Damian highlights OpenAI's ChatGPT for both research and teaching. He encourages its use for language processing tasks, such as literature searches, referencing, and data reformatting.

Power Automate – Although not considered full AI,
Power Automate can simplify and streamline workflows in academic research and administrative tasks.

Python – For more advanced AI tasks, python offers flexibility and powerful analytical capabilities for building and testing AI models.

Prompt Engineering – In his teaching, Damian encourages students to practice prompting techniques, effective prompts can significantly impact the quality and relevance of AI-generated responses.

Co-funded by the European Union

This project has been funded with support from the European Commission. The author is solely responsible for this publication (communication) and the Commission accepts no responsibility for any use may be made of the information contained therein. In compliance of the new GDPR framework, please note that the Partnership will only process your personal data in the sole interest and purpose of the project and without any prejudice to your rights.

ACCOUNTABILITY

Strategic AI Integration in Higher Education - Navigating Privacy, Bias, and Governance

SETTING

Danny Bielik is the President of the Digital Education Council, a Singapore-based organisation dedicated to helping universities and other higher education institutions understand and respond to the technological transformations brought about by AI. The Council works internationally, guiding institutions and governments on how AI is impacting education and the broader societal landscape. Danny draws on his extensive experience working with institutions around the globe, particularly focusing on AI's nascent role in higher education.

aileaders-project.eu
leaders
FACULTY OF MANAGEMENT University of Lodz
FEP
SCHOOL OF ECONOMICS AND MANAGEMENT UNIVERSITY OF PORTO
European E-Learning Institute
ACEEU ACCREDITATION COUNCIL FOR ENTREPRENEURIAL & ENGAGED UNIVERSITIES
upf. BARCE SCHOOL MANAGEMEN
Co-funded by the European Union
CLICK TO VIEW

> *"We're still at the tool stage of AI. We must look at the big picture and prepare for both transformational use cases and long-term societal change."*

CHALLENGE

Bielik identifies multiple challenges associated with the implementation of AI in higher education. Chief among these are data privacy, bias, accountability, and societal impact:

Data Privacy and Security: AI's reliance on vast amounts of data raises critical concerns around how student and institutional data are stored, managed, and potentially exploited. According to Bielik, privacy is the top concern among students, as evidenced by the Digital Education Council's global AI student survey, where 61% of respondents cited privacy and security as their primary worry. "Where are those data sets sitting? What happens with that data? We need to be able to assure students that their privacy is protected."

Bias in AI Models: AI systems often draw on data sets that favour developed regions or English-speaking populations, leading to biases in decision-making processes. This is particularly concerning in areas like university admissions, where biased algorithms could unfairly influence which students are admitted. Bielik stresses that institutions must have processes in place to evaluate and mitigate these biases.
"Bias is a problem, especially when AI is used in something like university admissions. We don't know what's in those data sets, and they often favour English-speaking, developed markets."

aileaders-project.eu
leaders
FACULTY OF MANAGEMENT University of Lodz
FEP SCHOOL OF ECONOMICS AND MANAGEMENT UNIVERSITY OF PORTO
celtech software innovations ltd
European E-Learning Institute
ACEEU ACCREDITATION COUNCIL FOR ENTREPRENEURIAL & ENGAGED UNIVERSITIES
upf. BARCELONA SCHOOL OF MANAGEMENT
Co-funded by the European Union
CLICK TO VIEW

Accountability and Governance: Who is responsible when an AI system makes an incorrect or biased decision? This question is central to Bielik's approach to AI governance. Institutions must have clear processes to monitor, evaluate, and respond to AI-related issues, ensuring that accountability lies not just with technology providers but with the institution itself. "Accountability isn't just about punishing people. It's about enabling processes that avoid land mines and achieve successful outcomes."

Regulatory Risks: As AI becomes more integrated into education, institutions face an evolving landscape of regulations that may differ across regions. These regulations could either stifle innovation or create new bureaucratic hurdles. The challenge lies in balancing the implementation of AI tools with adherence to data privacy laws like GDPR while ensuring that these tools remain effective. "Institutions need to have visibility from their vendors to manage data privacy. Universities must work closely with regulators and providers like Microsoft or Google to ensure compliance with data protection laws."

"Bias is a problem, especially when AI is used in something like university admissions. We don't know what's in those data sets, and they often favour English-speaking, developed markets."

CONCLUSION

Bielik stresses the importance of a strategic, structured approach to AI in education, focusing on long-term governance frameworks and processes that consider the specific needs of each institution. His key conclusions are:

Institutions must develop governance frameworks that are flexible and adaptable to their unique needs. These should include clear roles for accountability, monitoring tools, and strong communication mechanisms across faculty, students, and external stakeholders. Bielik advocates for gathering data from institutions to create benchmarks and guide future practices.

AI governance should prioritise transparency in how AI models function and how data is used. Students, faculty, and the wider community must be involved in discussions about AI's role, especially in terms of ethical usage, to avoid social backlash and maintain trust.

To avoid exacerbating global inequality, institutions must address uneven access to AI tools, particularly in developing countries. This will ensure that all students can benefit from AI, regardless of their geographic or socioeconomic background.

The societal transformation AI could bring is still in its early stages, but institutions must already begin to monitor how AI affects key human progress indicators, such as GDP, life expectancy, and equality. "We're still at the tool stage of AI. We must look at the big picture and prepare for both transformational use cases and long-term societal change."

"Accountability isn't just about punishing people. It's about enabling processes that avoid land mines and achieve successful outcomes."

DIVE DEEPER

<u>AI Governance Framework:</u> helps institutions establish clear accountability structures, communicate AI policies effectively, and gather data on AI use cases for benchmarking.

<u>Global AI Student Survey:</u> this survey highlights key student concerns about AI, with privacy and data security topping the list.

<u>Open AI Models:</u> language models can stimulate discussions and evaluate blind spots in academic work.

Co-funded by the European Union

This project has been funded with support from the European Commission. The author is solely responsible for this publication (communication) and the Commission accepts no responsibility for any use may be made of the information contained therein. In compliance of the new GDPR framework, please note that the Partnership will only process your personal data in the sole interest and purpose of the project and without any prejudice to your rights.

SOCIETAL IMPACT

Building Trust and Privacy in AI-Driven Civic Discourse

SETTING

Nicolas Gimenez is the co-founder and CTO of Zkorum, a platform that enables civic deliberation through technology, combining GovTech and social media to promote a bottom-up democracy. With a background in technology and a deep interest in privacy, Gimenez leads Zkorum's efforts to ensure that citizens can engage in public discourse without sacrificing privacy or trust.

Zkorum is a civic deliberation tool designed to foster healthy public debate and depolarise political conversations. The platform allows verified citizens to participate in discussions on various societal issues, using AI to classify and visualise opinions while preserving users' privacy. Zkorum incorporates zero-knowledge proofs (ZK) to ensure that users can verify their identity without exposing personal data.

The core of Zkorum's AI use lies in its recommendation algorithms and clustering models, which help group users based on shared opinions, showing the diversity of thoughts in the conversation and surfacing consensus across political divides. This innovative approach aims to reduce polarisation by identifying opinions that resonate across different ideological clusters.

> *"There's a lot of open washing going on... even Meta's so-called open-source AI is not truly open."*

aileaders-project.eu

Co-funded by the European Union. Views and opinions expressed are however those of the author or authors only and do not necessarily reflect those of the European Union or the Foundation for the Development of the Education System. Neither the European Union nor

leaders
FACULTY OF MANAGEMENT University of Lodz
FEP SCHOOL OF ECONOMICS AND MANAGEMENT UNIVERSITY OF PORTO
software innova
European E-Learning Institute
ACEEU ACCREDITATION COUNCIL FOR ENTREPRENEURIAL & ENGAGED UNIVERSITIES
upf. BARCE SCHOOL MANAGEME
Co-funded by the European Union

CLICK TO VIEW

CHALLENGE

aileaders-project.eu
leaders
FACULTY OF MANAGEMENT University of Lodz
FEP SCHOOL OF ECONOMICS AND MANAGEMENT UNIVERSITY OF PORTO
feltech software innovations ltd
European E-Learning Institute
ACEEU
upf. BARCELONA SCHOOL OF MANAGEMENT
CLICK TO VIEW

One of the critical challenges for Zkorum, as outlined by Gimenez, is the tension between open-source AI models and proprietary systems. Zkorum's commitment to open-source transparency ensures that the AI models they use are open for public scrutiny, allowing anyone to verify how the models function. However, the rise of open washing—where companies claim their AI is open but impose restrictions—makes it difficult for Zkorum to find reliable partners.

Gimenez remarked, "There's a lot of open washing going on... even Meta, training so-called open-source AI, is not truly open source. It's source-available, which isn't the same." Zkorum aims to ensure that both the AI models and data used are verifiable and open to public audit. However, this approach becomes increasingly difficult with more complex large language models (LLMs), which often have unpredictable outputs due to the vast number of parameters involved.

Zkorum's AI helps cluster users' opinions into coherent groups, providing insights into political or social divisions. However, as they explore further AI applications, such as natural language querying of discussion results or assistance with writing comments, they face the challenge of ensuring that AI does not unduly influence or censor users' voices. Maintaining the balance between AI moderation and free expression is a delicate task. "Rephrasing comments using AI is very sensitive... the challenge is not making the user feel like they're being censored," Gimenez explained. This is a particular concern when dealing with emotionally charged topics, where users expect their voice to be preserved, even if AI aids in formatting or clarity.

Privacy is at the heart of Zkorum's mission. The platform uses zero-knowledge proofs to verify users' identities without exposing personal information. However, as Zkorum explores more advanced AI use cases—such as fact-checking or querying data sets—Gimenez is mindful of how they handle user data. He stressed that AI systems must be run on local devices where possible to avoid compromising user privacy.

Communicating the transparency and accountability of AI systems to users is a complex issue, particularly when even tech-savvy users struggle to fully understand open-source principles. Although Zkorum uses open-source AI to build trust, Gimenez acknowledges that many users will never have the technical knowledge to verify this themselves.
"Even with open source, users still need to trust that the model running is actually the open-source one," said Gimenez, underscoring the importance of building trust through transparency.

"Rephrasing comments using AI is very sensitive... the challenge is not making the user feel like they're being censored."

CONCLUSION

Zkorum's use of AI is framed by a strong commitment to open-source transparency, privacy, and the promotion of healthy civic discourse. Gimenez emphasises that while AI—particularly LLMs—holds great potential, it must be deployed carefully to avoid undermining users' voices or privacy.

The platform's AI is not an end in itself but a tool for enabling collective intelligence, where deliberation is enriched by thoughtful clustering of opinions and visualisation of consensus across political divides. Gimenez's focus on open-source AI ensures that both the algorithms and data remain verifiable, but he is also pragmatic about the challenges of ensuring trust in such systems.

While Zkorum's AI models are used primarily for clustering and classification, the platform is poised to expand its AI capabilities with natural language querying, fact-checking, and comment assistance. Each of these use cases will require further refinements to balance the role of AI in augmenting public debate without eroding trust or privacy.

"Even with open source, users still need to trust that the model running is actually the open-source one."

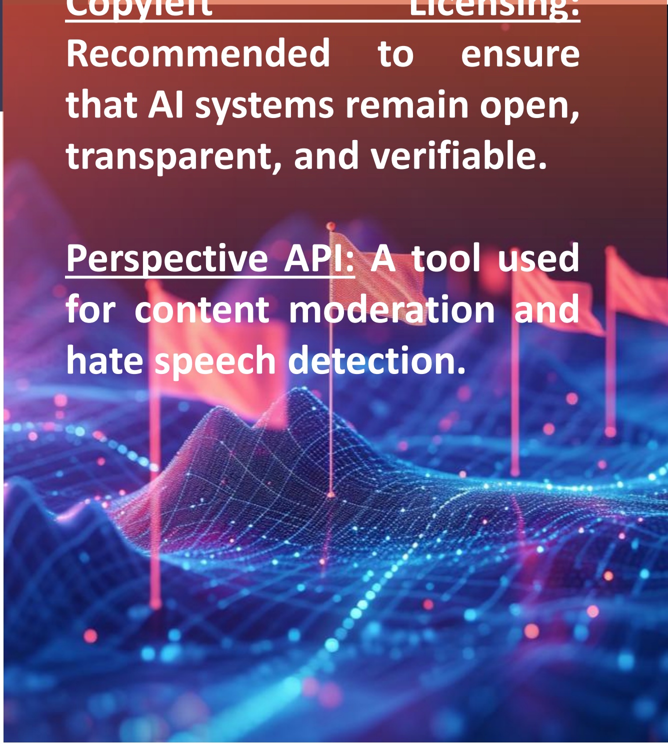
DIVE DEEPER

Pol.is: Used for clustering and analysing civic discussions, grouping participants based on shared opinions and behaviours.

Zero-Knowledge Proofs (ZK): Employed for privacy-preserving identity verification.

Helix.ml: An open-source platform for fine-tuning AI models and understanding how LLMs work.

Copyleft Licensing: Recommended to ensure that AI systems remain open, transparent, and verifiable.



Perspective API: A tool used for content moderation and hate speech detection.

Co-funded by the European Union

This project has been funded with support from the European Commission. The author is solely responsible for this publication (communication) and the Commission accepts no responsibility for any use may be made of the information contained therein. In compliance of the new GDPR framework, please note that the Partnership will only process your personal data in the sole interest and purpose of the project and without any prejudice to your rights.

FAIRNESS

Building Trustworthy AI: Navigating Ethics, Equity, and Innovation in Global Healthcare

SPOTLAB

SETTING

Miguel Luengo-Oroz is the Co-Founder and CEO of Spotlab.ai, a spin-off from the Technical University of Madrid. With a background in engineering, cognitive science, and a PhD focused on AI for medical image processing, Miguel's career spans over two decades in both science and global policy. Notably, he served as the Chief Data Scientist for the United Nations, working on innovation projects for humanitarian response and global health, as well as crafting AI-related policies. Spotlab.ai, founded in 2018, is dedicated to developing trustworthy AI for diagnostics and clinical research, making cutting-edge technology accessible to both high-resource settings and low-income regions worldwide.

aileaders-project.eu
leaders
FACULTY OF MANAGEMENT University of Lodz
FEP SCHOOL OF ECONOMICS AND MANAGEMENT UNIVERSITY OF PORTO
software innova
European E-Learning Institute
ACEEU ACCREDITATION COUNCIL FOR ENTREPRENEURIAL & ENGAGED UNIVERSITIES
upf. BARCE SCHOOL MANAGEMEN
Co-funded by the European Union
CLICK TO VIEW

> *"Mitigating bias is step 0... We need to know what we don't know. Mapping that universal bias is critical."*

CHALLENGE

One of the core challenges for Spotlab.ai involves mitigating bias in AI systems—a critical concern in ensuring equitable healthcare outcomes. Miguel emphasised that managing bias is a "step 0" for Spotlab.ai, as they map potential biases from the initial stages of model development. This involves evaluating biases in data sources, such as differences between medical samples from diverse regions or economic backgrounds, and addressing variability in data acquisition, whether from high-end scanners or mobile phones. Ensuring fairness requires careful handling of how data is labelled and analysed, incorporating diverse perspectives to minimise observer bias. By embedding these processes into every stage, Spotlab.ai strives to create diagnostic tools that work for "everyone, everywhere."

Ethical and regulatory compliance is another complex area. Miguel outlined Spotlab.ai's adherence to six core AI principles identified by the World Health Organisation (WHO), including protecting human autonomy, promoting well-being, ensuring transparency, accountability, equity, and sustainability.

aileaders-project.eu

leaders

FACULTY OF MANAGEMENT University of Lodz

FEP SCHOOL OF ECONOMICS AND MANAGEMENT UNIVERSITY OF PORTO

celtech software innovations ltd

European E-Learning Institute

ACEEU ACCREDITATION COUNCIL FOR ENTREPRENEURIAL & ENGAGED UNIVERSITIES

upf. BARCELONA SCHOOL OF MANAGEMENT

Co-funded by the European Union. Views and opinions expressed are however those of the author or authors only and do not necessarily reflect those of the European Union or the Foundation for the Development of the Education System. Neither the European Union nor the entity providing the grant can be held responsible for them.

Co-funded by the European Union

CLICK TO VIEW

Spotlab.ai is navigating a field regulated by multiple overlapping frameworks, from medical device standards to privacy laws like GDPR and cybersecurity requirements. This has necessitated a shift in organisational mindset, from traditional research practices to a model that prioritises long-term safety and ethical accountability.

A broader societal challenge involves expanding access to high-quality healthcare diagnostics. Spotlab.ai addresses this by applying the same AI technology in Europe's leading hospitals to low-resource settings in countries like Ethiopia and regions in Latin America. This approach aims to close the diagnostic gap, ensuring that AI-driven solutions serve diverse global populations, including those affected by neglected tropical diseases like leishmaniasis and Chagas.

CONCLUSION

Spotlab.ai's AI-driven solutions have achieved significant impact by reducing diagnostic errors and speeding up access to appropriate treatments. Their focus on fairness, transparency, and regulatory compliance has built trust among stakeholders, from clinicians in high-resource settings to healthcare workers in low-income regions. The decision to deploy diagnostics on smartphones exemplifies the company's dedication to accessibility and equity, allowing AI models to operate in areas with limited connectivity.

Spotlab.ai's experiences highlight the importance of having a "compass over maps"—a clear direction without rigid plans—allowing them to adapt as regulations and technologies evolve. Miguel's leadership has underscored the need to bridge the language of AI with the subject matter expertise in healthcare, ensuring that both clinicians and AI developers work towards common goals. The company's emphasis on societal impact aligns with Sustainable Development Goal 3 (SDG-3), promoting good health and well-being on a global scale.

"Compass over maps. I know where I'm going, but I don't have a map because many things are changing constantly and we need to adapt ."

DIVE DEEPER

World Health Organisation's AI Principles: Spotlab.ai follows WHO's guidelines, which emphasise ethical AI development, including autonomy, transparency, and equity.

European Union Regulations: Compliance with multiple regulatory frameworks, including GDPR for data privacy, ISO standards for data security, medical devices, specific sectoral regulations , and the evolving AI Act, is a key part of Spotlab.ai's operations.

SDG-3: The company's commitment to reducing healthcare disparities aligns with this goal, targeting better health and well-being globally, particularly in underserved regions.

Co-funded by the European Union

This project has been funded with support from the European Commission. The author is solely responsible for this publication (communication) and the Commission accepts no responsibility for any use may be made of the information contained therein. In compliance of the new GDPR framework, please note that the Partnership will only process your personal data in the sole interest and purpose of the project and without any prejudice to your rights.

DIGITAL TOOLBOX

www.aileaders-project.eu

Co-funded by the European Union

Co-funded by the European Union. Views and opinions expressed are however those of the author or authors only and do not necessarily reflect those of the European Union or the Foundation for the Development of the Education System. Neither the European Union nor the entity providing the grant can be held responsible for them.

Glossary of AI Terms

Algorithm: Procedure of steps or instructions a computer follows that can be used by machine learning systems to ingest data and make predictions.

Artificial Intelligence (AI): Simulation of human intelligence by machines or computers.

AI Ethics: Principles that govern AI's behavior in terms of human values.

Bias: Prejudiced results or outputs produced by algorithms or impartial data.

Big Data: Extremely large information sets that may be analyzed computationally to reveal patterns, trends, and associations, especially relating to human behavior and interactions.

Bounded input-bounded output (BIBO) stability: A system that indicates if every bounded input leads to a bounded output.

Chatbot: AI-powered tool intended to communicate in a conversational manner.

Copyleft: A form of licensing for Open-Source Code usage, ensuring a program, code, or other similar work, as well as subsequent versions or modifications, are free of cost. Helps ensure that AI systems remain open, transparent, and verifiable.

Cutoff date: The end date for an AI model's data input.

Data mining: Process of searching for patterns within a large set of data to extract specific information.

Data validation: Process of checking data quality before using it for AI models.

Deep learning: machine learning that imitates how humans learn new information.

Emergent Behavior (aka emergence): Unpredictable or unintended capabilities shown by AI systems

EU AI Act: Regulation for AI deployment within the European Union, focused on a risk-based approach to AI that classifies systems according to their potential impact on fundamental rights, including, among others, data privacy rights.

Fairness: Creation of algorithms and systems that make decisions impartially, equitably, and justly across all individuals and groups, especially those considered sensitive or protected, such as different ethnicities, genders, or disabilities.

Garbage in, garbage out (GIGO): Computer science concept that notes the quality of the output depends on the quality of the input.

General Data Protection Regulation (GDPR): European regulation that sets guidelines for the collection and processing of personal information from individuals who live in and outside of the European Union (EU).

Generative AI: AI technology that creates content from learned patterns in data.

Generative Pre-Trained Transformer (GPT): AI algorithm.

Guardrails: Rules or restrictions applied to AI models to direct data use.

Hyperparameter: Configuration variables that data scientists set ahead of time to manage the training process of a machine learning model.

Knowledge Engineering: AI field that tries to emulate (expert) human knowledge.

Large Language Model (LLM): AI algorithm that understands, summarizes, generates, and predicts new content.

Machine Learning: Use and development of computer systems that can learn and adapt without following explicit instructions, by using algorithms and statistical models to analyze and draw inferences from patterns in data.

Moats: The ability for a product or company to maintain a competitive advantage and fend off competition to maintain profitability into the future.

Multimodal AI: Machine learning models capable of processing and integrating information from multiple modalities or types of data.

Natural Language Generation (NLG): Application of AI models to create written or spoken narratives that make sense to humans.

Glossary of AI Terms

Natural Language Processing (NLP): Branch of AI that enables computers to comprehend, generate, and manipulate human language.

Open Washing: When AI models are claimed to be open access but have user restrictions.

Optical Character Recognition (OCR): Technology that uses automated data extraction to quickly convert images of text into a machine-readable format.

Overfitting: Undesirable machine learning behavior that occurs when the machine learning model gives accurate predictions for training data but not for new data.

Parameter: A value that is used to control the operation of a function or that is used by a function to compute one or more outputs.

Pattern Recognition: Ability of machines to identify patterns in data, and then use those patterns to make decisions or predictions using computer algorithms.

Predictive AI: Use of machine learning to identify patterns in past events and make predictions about future events.

Prompt: Text or symbols used to represent the system's readiness to perform the next command.

Real-Time AI Feedback: Immediate responses or actionable insights provided by an artificial intelligence (AI) system based on live data inputs.

Structured Data: Defined and searchable data.

Synthetic Data: Artificial data that is generated from original data and a model that is trained to reproduce the characteristics and structure of the original data.

Training Data: Large dataset used to train machine learning (ML) models to process information and accurately predict outcomes.

Transfer Learning: Machine learning technique in which knowledge gained through one task or dataset is used to improve model performance on another related task and/or different dataset.

Unstructured Data: information that does not have a fixed format or structure that makes it difficult to organize and analyze.

Unsupervised Learning: A type of machine learning that learns from data without human supervision. Unlike supervised learning, unsupervised machine learning models are given unlabeled data and allowed to discover patterns and insights without any explicit guidance or instruction.

Company Bios

Digital Education Council (DEC) is an organization dedicated to advancing education through modern digital tools and strategies. By fostering collaboration among educators, innovators, and policymakers, the DEC aims to equip learners with the skills necessary to thrive in a technology-driven world, promoting equitable access to quality educational resources and lifelong learning opportunities.

ZKorum is a startup dedicated to building open-source protocols and softwares to rehumanize and depolarize the online social landscape. It's main project is Agora by ZKorum,, open-source under the AGPLv3 licence, it is an eDemocracy tool designed to empower individuals to engage in free civil discourse to find consensus, paving the way for more resilient societies. With Agora any interested party to audit the moderation of content — ensuring that moderators are accountable for every moderation decision. The solution will be based on Nostr, a P2P protocol to broadcast the proof (metadata) of each post so that when a social media platform, centralised or federated, removes a post, anyone can independently use unforgeable cryptographic proofs to hold the platform accountable. This solution will make users less dependent on trusting platforms, a principle known as "trust-minimization."

Spotlab.ai is a health technology company that harnesses artificial intelligence and telemedicine to improve medical diagnostics and research. By combining advanced imaging, remote collaboration tools, and deep-learning capabilities, Spotlab.ai helps healthcare professionals and organizations access, analyze, and share critical data more efficiently, ultimately enhancing patient care and fostering innovation in global healthcare delivery.

moveUP is a Brussels-based digital health company offering AI-powered rehabilitation and remote patient monitoring across orthopaedics, cardiology, oncology, and other clinical areas. Its platform supports personalised care by combining predictive analytics, real-time data collection, and automated decision support for healthcare professionals. Patients benefit from tailored education and support programmes, while healthcare providers gain tools to streamline workflows and improve outcomes. Operating across several countries, moveUP's approach has led to high patient satisfaction and reduced unplanned consultations. The company recently expanded its AI capabilities through the acquisition of DeepStructure.ai, reinforcing its role as a leader in data-driven, patient-centred care.

Tool Zero is a Canadian software company simplifying carbon accounting for SMEs through AI automation. Its core product, Emissions AI, extracts data from invoices to calculate Scope 1–3 emissions, matched with verified emission factors across 49 regions. Designed for businesses without in-house sustainability teams, Tool Zero's platform supports regulatory compliance, emissions reporting, and funding access. The tool helps visualise carbon hotspots, supporting informed decision-making and sustainability certification. Based in Kitchener, Ontario, Tool Zero partners with clean tech incubators to promote affordable, credible, and automated GHG tracking for small businesses navigating the green transition.

Company Bios

Feedzai is a fintech company that provides AI-based solutions for financial fraud detection, using purpose-built AI to detect and prevent fraud in real-time. While Feedzai's core business is fraud detection, developing and deploying AI systems in financial institutions, the company also helps institutions achieve compliance and prevent illicit activities by uncovering money laundering and organized crime. Its end-to-end platform covers the entire financial crime lifecycle, from account opening to fraud prevention and AML compliance.

LTP Labs is a consulting firm that combines Advanced Analytics and AI expertise with a strong business acumen. They are committed to transforming companies and empowering every decision with Analytics, and specialize in identifying, igniting and scaling high-impact use cases that drive significant value using Advanced Analytics and AI. LPT Labs has conducted projects internationally across the globe, and also across multiple sectors, from retail to healthcare.

Cobiro is a Danish MarTech company offering AI-powered tools that automate and optimise digital advertising for SMEs across Europe. By simplifying complex marketing processes, Cobiro enables businesses with limited technical expertise to run effective, data-driven campaigns. Committed to fairness and transparency, the company continuously monitors its algorithms to reduce bias and ensure explainability. CEO Jesper Valentin Holm advocates for responsible AI use, combining automation with human oversight. Cobiro complies with GDPR and prioritises ethical data use helping SMEs compete in digital markets while making AI understandable and accessible.

Plytix is a Danish company offering AI-powered product information management (PIM) software that helps SMEs optimise and automate their multichannel retail operations. Serving clients globally, Plytix uses AI to generate high-quality, platform-specific product content at scale, improving efficiency and consistency across e-commerce platforms. The company places strong emphasis on fairness, transparency, and privacy, with systems designed to detect bias, respond to user feedback, and comply with GDPR. CEO Morten Poulsen believes in using AI to solve specific, real-world problems—streamlining data workflows while maintaining ethical and accountable development practices.

AI Heroes is a European company delivering AI-powered solutions to streamline operations, automate tasks, and enhance innovation across sectors. From customer support automation to fall detection in elderly care, their tools improve efficiency while upholding privacy and ethical standards. Led by Roan Van Der Sluis, the company applies both AI-first and AI-second approaches to balance automation with human creativity. AI Heroes actively addresses bias, ensures GDPR-compliant data practices, and engages in ethical oversight throughout project development—demonstrating how AI can empower people, not replace them.

leaders

Follow Our Journey

n

www.aileaders-project.eu

Co-funded by the European Union

Co-funded by the European Union. Views and opinions expressed are however those of the author or authors only and do not necessarily reflect those of the European Union or the Foundation for the Development of the Education System. Neither the European Union nor the entity providing the grant can be held responsible for them.