Game-Based Learning as a Tool for Teaching Ethical AI to Youth: Insights from the CHARLIE Project

Nidhi[1], M. Begoña Arenas[2], Andreia Morgado[3], Adrian Solomon[4], Maria Moreira[5], Rute Ferreira[5] and Elise Raittila[6]
[1]Aarhus Universitet, Denmark
[2]UNIR Universidad Internacional de la Rioja, Spain
[3]IMPACTsci, Portugal
[4]HELIXCONNECT EUROPE S.R.L, Romania
[5]SQ e-learning, S.A., Portugal
[6]Vaasan Ammattikorkeakoulu; University of Applied Sciences; Finland

nidhi@btech.au.dk
begona.arenas@unir.net
andreia.morgado@impactsci.org
andrian.solomon@helix-connect.com
maria.moreira@isqe.com
rute.ferreira@isqe.com
elise.raittila@muova.fi

Abstract: The pervasive influence of Artificial Intelligence (AI) and Machine Learning (ML) technologies in everyday life has underscored the critical need for addressing inherent biases in big data. The CHARLIE project, an "*ERASMUS+ KA2"* initiative involving six partners across five European countries, is a beacon of innovation in addressing the critical need for algorithmic biases and ethical education in AI and ML. Our project objectives centre on enhancing the capacity of Higher Education (HE) institutions to deliver socially responsible and ethically informed tech education, particularly in AI and ML domains. By integrating digital and engaging pedagogical strategies, we aim to increase tech students' social and ethical competencies and equip educators with effective tools for teaching these critical topics. CHARLIE focuses on cultivating an ethical, human-centered perspective in technology education, specifically targeting the youth demographic. One of the practical outputs from CHARLIE is the development of a competency matrix and the 'Algorithmic Bias' course. This course, designed with a blended learning approach and complemented by a 'Digital Game' to assess learning outcomes, is a testament to the project's real-world applicability. Additionally, the project focuses on creating effective linkages between HE, Adult Education (AE), and Youth sectors to facilitate widespread adoption and recognition of ethical AI curricula. This method is chosen for its potential to engage young learners in complex subjects by integrating learning with interactive, scenario-based gameplay that reflects real-world dilemmas. The games complement theoretical knowledge with practical decision-making exercises that promote critical thinking and ethical reasoning. Our paper will outline the development process of our digital games, discuss the pedagogical frameworks that underpin them, and share preliminary results from their implementation across various educational settings in Europe. We will also explore the broader implications of our work for integrating game-based learning in ethical tech education and its potential to transform educational practices in Higher Education, Adult Education, and beyond.

Keywords: Game-Based Learning, Ethical AI, Youth Education, Digital Games, Bias in AI, Technology Education.

1. Introduction

1.1 Background

Artificial Intelligence (AI) and Machine Learning (ML) are deeply integrated into our daily lives. These technologies influence numerous aspects of society, from facial recognition systems used in public security to algorithms determining the advertisements we see on social media. However, despite their mathematical foundations, AI and ML systems are not infallible; they often perpetuate biases in their training data. Addressing these biases is crucial, as large tech companies' widespread adoption of AI and ML—primarily driven by profit motives—raises significant ethical concerns.

Human bias, extensively studied by social scientists, emerges from implicit associations and unconscious prejudices. These biases can lead to adverse outcomes, from perpetuating stereotypes to reinforcing systemic discrimination. AI and ML systems, designed to make predictions based on existing data, inherently reflect and amplify these biases. Without an inherent mechanism for ethical decision-making, these technologies often perpetuate and exacerbate existing societal issues.

The sophistication of AI technology presents significant ethical challenges, necessitating stringent risk management procedures to ensure quality delivery. It is crucial to maintain flexibility that respects human values as AI advances, promoting sustainable innovations (Nguyen et al., 2023). In response to this need, UNESCO established global standards for AI ethics, endorsed by its 193 member countries on November 25, 2021. In the document presented, they acknowledge the "profound and dynamic" impact of AI while also highlighting the associated risks to cultural, social, and ecological diversity (UNESCO, 2021).

Incorporating AI tools in the educational sector has become a driving force in transforming learning experiences, encouraging innovation, and equipping individuals for the digital era. Despite this encouraging prospect of AI in education, there needs to be more literature addressing the ethical challenges involved (Abulibdeh et al., 2024). AI chatbots, like ChatGPT, have attracted significant attention and can alter various educational facets dramatically. Nevertheless, their implementation brings up ethical dilemmas, necessitates the redesign of curricula, calls for strategies to facilitate ongoing learning, and requires synchronization with industry norms. Additionally, although it is widely recognized that education must align with the changing demands of Industry 4.0, there still needs to be more research on how academic institutions can effectively partner with industries to close the gap between education and industry requirements. This encompasses ethical considerations related to AI-generated content, the development of students' critical thinking skills, the competencies of both teachers and students, concerns about data privacy, and potential biases (Abulibdeh et al., 2024).

Regarding generative AI, most created content is favourable or potentially beneficial to groups. However, a significant peril exists wherein these systems might inadvertently propagate profoundly harmful content. These emerging generative AI frameworks can propagate misinformation, distortions, prejudices, and hazardous content. It is crucial to note that with the advancement of generative AI tools, producing such content becomes cost-effective and more accessible. Furthermore, pre-existing harmful content could serve as a blueprint for generating even more harmful material (Fedele et al., 2024).

Trustworthy AI, in this sense, is the capability of an AI-driven system to optimize its advantages while mitigating and forestalling potential risks. This principle was initially outlined in the Ethical Guidelines for Trustworthy AI (EC, 2019), which outlined its three fundamental characteristics: legality, ethicality, and resilience.

1.2 Problem Statement

The European Commission (EC) has articulated a vision for "ethical, secure, and cutting-edge AI made in Europe," emphasizing the importance of human-centered AI development (EC, 2019). AI can potentially transform society positively by promoting gender equality, tackling climate change, enhancing health and mobility, and supporting sustainable development goals. However, to realize these benefits, AI systems must be developed and implemented in ways that prioritize human welfare, freedom, and the common good.

AI systems present significant opportunities and risks. To maximize their benefits and mitigate their risks, AI development must be guided by ethical principles (Khan et al., 2022). Trustworthy AI, which incorporates ethical considerations into its design and implementation, can provide a competitive advantage to its producers and foster public trust in sociotechnical systems.

Higher Education (HE), Adult Education (AE), and Youth sectors require innovative curricula addressing AI and ML's ethical implications. Educational programs that equip learners with the knowledge and skills to develop and deploy AI systems ethically are critical (Slimi et al., 2023). The Digital Education Action Plan highlights the necessity of addressing ethical challenges in AI and data usage within education and training frameworks.

The first identified need is societal: the need for a fair society where embedded technology does not replicate societal biases and prejudice against minorities. The rapid advancement of new technologies brings inherent risks and challenges, necessitating innovations such as AI in educational contexts that are ethical, inclusive, and equitable. Psychologists identify approximately 180 cognitive biases, some of which may influence AI algorithms' design, including racial bias, gender prejudice, recruiting inequity, and age discrimination (Schwartz et al., 2022).

Individuals suffering from biases, such as machismo, sexism, racism, xenophobia, and other forms of discrimination (e.g., minorities and disadvantaged groups), urgently need technology to help level opportunities rather than perpetuate existing biases. For instance, AI could exacerbate challenges such as job interview disparities and biased criminal identification, disproportionately affecting non-white individuals.

While many universities offer ethics courses within social sciences, these often need an interdisciplinary approach. An interdisciplinary approach is essential for students to graduate with comprehensive knowledge

and skills applicable to their careers (Kayyali, 2024). Higher education teachers also need enhanced skills to provide effective online learning opportunities, moving beyond essential tools like PowerPoint presentations in Zoom or Microsoft Teams to true eLearning methodologies. CHARLIE will promote DigComp awareness among educators and recommend methodologies to identify, understand, and avoid biases in teaching, linking bias with AI ethics and producing state-of-the-art educational materials. Additionally, HE administrators responsible for IT management of eLearning (e.g., managing Moodle) need clear instructions on adopting eLearning methodologies and courses from both technical and pedagogical perspectives.

Adult education providers and universities need a more collaborative approach to supporting adult upskilling and reskilling (Lang, 2023), especially in AI, which offers significant employment opportunities for disadvantaged adults. However, adult education providers often need more expertise to develop engaging eLearning tailored to adult learners' needs.

Youth, particularly those aged 12-18, young women, and those from disadvantaged backgrounds, must be aware of potential futures in emerging technologies and AI. There is a pressing need for awareness initiatives targeting this group and developing engaging and fun Open Educational Resources (OERs) for youth organizations.

1.3 Project Overview

The CHARLIE project, an *"ERASMUS+ KA2 (2022-1-ES01-KA220-HED-C461966C)"* initiative, seeks to confront biases in big data utilized by AI and ML technologies. The project emphasizes the critical need for a comprehensive and ethical approach to Technology Education (techEd). CHARLIE aims to revolutionize tech education in Europe's HE, AE, and youth sectors by raising awareness of the negative impacts of a lack of ethical consideration in AI and ML. Coordinated by Universitat de les Illes Balears (UIB) in Spain, it involves partners from ISQ e-learning (ISQe) in Portugal, Innovation Training Center (ITC) in Spain, Helixconnect Europe (HELIX) in Romania, Aarhus Universitet (AU) in Denmark, and OY Vaasan Ammattikorkeakoulu—Vasa Yrkeshögskola AB (VAMK) in Finland, with IMPACTsci—Portugal as an associated partner.

CHARLIE's objectives are multifaceted. It seeks to augment the capacity of Higher Education institutions to deliver online learning opportunities that not only cater to societal needs but are also tailored to the diverse learning requirements of students. Charlie aims to elevate tech students' social and ethical competencies, enabling them to engage positively, critically, and ethically with AI and ML technologies. Furthermore, it aims to equip teachers and professors with digital and engaging approaches to teaching these topics, focusing on online settings effectively. The project also endeavours to foster collaboration between Higher Education, Adult Education, and youth organizations in AI ethics education. It aims to transfer academic courses on AI biases to the Adult Education and Vocational Education and Training (VET) sectors. Lastly, increasing societal awareness of the ethical implications of AI and ML is a critical goal.

The CHARLIE project pioneers several innovative outputs aimed at reshaping ethical AI education. Some of the key project outputs include:

- Competency matrices for "Algorithmic Bias" (EQF6), "Ethical AI microcredential" (EQF4), and the Serious Game (EQF2).
- A blended learning "Algorithmic Bias" course with a digital serious game for assessment.
- A toolkit for synchronous sessions and guidelines for university administrators.
- A self-paced, online "Ethical AI Microcredential" for adult learners.
- A digital serious game for youth aged 12-18, targeting disadvantaged groups and promoting gender representation in STEM.
- Policy recommendations to recognize the microcredential for facilitating access to HE courses in technological fields.

1.4 Purpose of the Paper

This paper aims to analyse the effectiveness of the game-based learning approach in achieving the project's educational goals by looking at various structural pedagogical design elements of ethical AI frameworks in higher education. In particular, the paper will outline the development process of our digital games, discuss the pedagogical frameworks that underpin them, and share preliminary results from their implementation across various educational settings in Europe. The paper will also explore the broader implications of the named project for integrating game-based learning in ethical tech education and its potential to transform educational practices in Higher Education, Adult Education, and beyond.

2. Methodology

The project is structured into five Work Packages (WPs):

- **WP1—Project Management and Evaluation:** Coordinates and supports all partners, ensuring timely goal achievement and compliance with obligations.
- **WP2 - Competency Matrices for Algorithmic Bias:** Establishes learning goals for HE students, AE learners, and youth through the "Algorithmic Bias" course, "Ethical AI" micro-credential, and serious game.
- **WP3 - Algorithmic Bias Toolkit for HE:** Develops OERs and tools to help HE institutions adopt ethical AI and ML curricula.
- **WP4 - Transfer of Products to AE and Youth:** Upskills adults and youth, raises awareness about ethical AI, and provides tailored digital OERs.
- **WP5 - Communication and Project Promotion:** Disseminates project activities and results to a broad audience, emphasizing HE institutions' role in addressing AI biases and promoting awareness.

The CHARLIE project partners have designed WP2 Competency Matrices for learning programs aligned with EQF levels: EQF6 for Higher Education, EQF4 for Adult Education, and EQF2 for youth. They are now working to achieve the project's expected outcomes.

2.1 Target Groups

As to our targets, the CHARLIE project seeks to foster a comprehensive and inclusive approach to AI education, addressing the needs and challenges of various demographic groups across HE and AE institutions and the youth sector:

Higher Education (HE) Institutions: The project aims to enhance the capacity of HE institutions by developing and integrating ethical approaches to AI and ML into their curricula. This includes creating competency matrices, courses, toolkits, and guidelines for HE students, teachers, and administrators. The goal is to equip these institutions with the resources to address and mitigate algorithmic biases effectively.

Adult Education (AE) Institutions: The project also focuses on adult learners by providing specialized resources and learning pathways tailored to their needs. This includes the development of an Ethical AI micro-credential and competency matrices aligned with EQF4. The aim is to upskill adults in ethical AI practices, enabling them to engage with and understand the implications of AI/ML in their professional and personal lives.

Youth: The CHARLIE project targets young people, particularly those aged 12-18, focusing on young women and youth from socially disadvantaged backgrounds. The project aims to raise awareness and interest in ethical AI by developing a digital serious game and competency matrices aligned with EQF2. By engaging youth in a fun and interactive manner, the project seeks to inspire the next generation to consider careers in ethical AI and understand the societal impacts of technology.

2.2 WP3 - Algorithmic Bias Toolkit for HE

In this work package, the partnership focuses on developing comprehensive resources and tools to help HE institutions integrate ethical considerations into their AI and ML curricula. CHARLIE employs a collaborative and co-creation approach to develop its educational resources focused on six core Competence Units (CUs): Algorithm Models and Limitations, Data Fairness and Bias in AI, AI Privacy and Convenience, AI Ethics in Practice, and Case Studies and Projects. The partnership has designed an effective collaborative and co-creation approach, as illustrated below in Figure 1.

Course Architecture Development	ISQe develops the course architecture and templates and then presents them for partner approval.
Template Completion	UIB, AU, and VAMK complete templates with learning outcomes based on the five CUs.
Peer Review and Multimedia Development	ISQe handles multimedia development, validation tests, and final adjustments.
Template Finalization	Partners fill templates; ISQe finalizes multimedia and delivers the toolkit, including PowerPoint templates.
Webinar Guidelines and Organization	HELIX provides guidelines and finalizes the webinar agenda; partners organize outreach events.
Translation and Delivery	ISQe distributes translation templates; partners complete translations and deliver results in all relevant languages.
Pilot Course Implementation	UIB, AU, and VAMK implement the "Algorithmic Bias" pilot course with 90 participants total.
Quality Check	ITC and all partners conduct a quality check to ensure standards are met.

Figure 1: WP3 Workflow

2.3 WP4 - Transfer of Products to AE and Youth

WP4 promotes ethical AI practices across various demographics, particularly in the AE and youth sectors. Led by ITC and supported by ISQe, this phase includes developing educational resources like an Ethical AI Microcredential and a Digital Serious Game. The aim is to equip learners and stakeholders with essential knowledge and skills to foster responsible AI usage. Additionally, efforts are directed towards creating a toolkit to support adults and youth in Ethical AI upskilling. The project aims to equip learners and stakeholders with essential knowledge and skills through these endeavours, fostering a responsible approach to AI technologies.

The CHARLIE project is structured around six core competence units: CU1 focuses on understanding algorithmic bias, CU2 addresses the principle of non-maleficence, CU3 emphasizes accountability, CU4 highlights the importance of transparency, CU5 covers human rights and fairness, and CU6 provides a practical approach to AI ethics. These competencies are integrated into the Digital Serious Game, designed as an escape room with four thematic rooms. These rooms cover the introduction to AI and potential inequalities, basic mechanics of AI systems and algorithmic biases, identifying and addressing algorithmic bias, and ethical considerations in AI, including policy agendas and global impact. Figure 2 illustrates a mapping diagram of how the six competence units are integrated into the four rooms of the Digital Serious Game.

Room 1: Introduction to AI and Potential Inequalities	CU1: What is Algorithmic Bias?
Room 2: Basic Mechanics of AI Systems and Algorithmic Biases	CU2: Non-maleficence CU4: Transparency
Room 3: Identifying and Addressing Algorithmic Bias	CU3: Accountability CU6: AI Ethics, a Practical Approach
Room 4: Ethical Considerations in AI: Policy Agendas and Global Impact	CU5: Human Rights and Fairness

Figure 2: Mapping Diagram for Competence Units to Game Rooms

The game combines EntreComp, DigComp 2.0, and GrenComp competencies to ensure comprehensive learning, engaging students through interactive challenges and feedback. Figure 3 summarizes the workflow of the WP4.

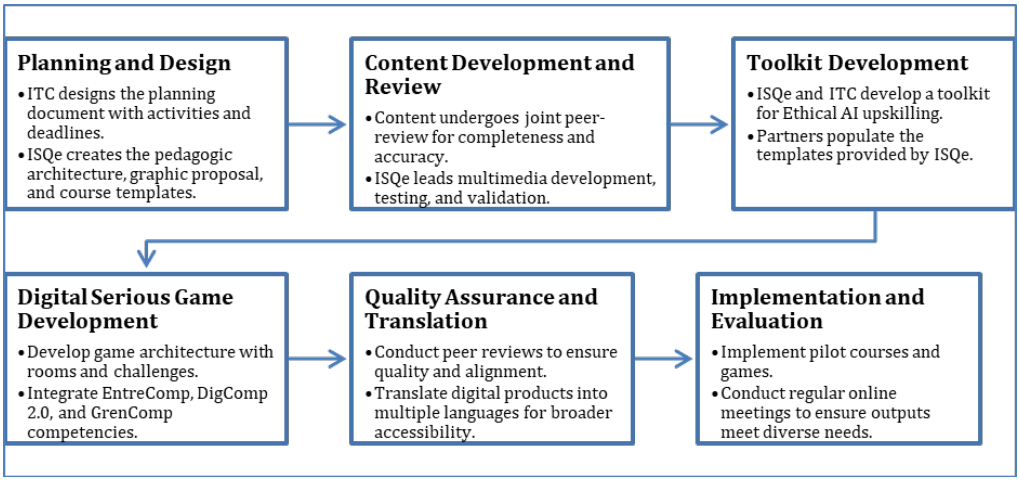

Figure 3: WP4 Workflow

3. Educational Framework and Implementation

To create a more engaging and interactive learning experience, the “Algorithmic Bias” course was developed as an integrated blended learning course with a digital roulette game for formative assessment. The course includes flipbooks, quizzes, project-based exercises, expert videos, concept maps, and a support toolkit for trainers. Thorough research and iterative testing ensure that the materials meet learners’ needs and that serious games create dynamic learning environments that enhance engagement and practical understanding. This approach helps to bridge theory with practice, thereby equipping learners to tackle algorithmic bias effectively.

3.1 Course Development

To address the Algorithmic Foundations and Ethics in AI, an instructional design team has set themselves on making a comprehensive course that uses blended learning and builds on it with a digital roulette game in addition to multiple choices questions and scoring as a final formative assessment exercise, to address Algorithmic Foundations and Ethics in AI. This innovative course structure aims to provide an engaging, interactive, and effective learning experience. The pedagogical architecture is the starting point of creating the "Algorithmic Bias" course. Clear learning objectives are established to ensure learners understand algorithmic bias, its implications, and mitigation strategies. The curriculum incorporates various eLearning components such as interactive screens, flipbooks, multiple choice quizzes, project-based exercises, short expert videos, interactive concept maps, and the roulette game.

Such pedagogical architecture also entails developing a range of support tools for trainers, including:

- **Support PowerPoint Presentations:** These will help deliver key concepts consistently by providing structured content during synchronous sessions.
- **Handbook:** This detailed guide covers course objectives for trainers, content summary teaching strategies, and troubleshooting tips.
- **Multiple Choice Quizzes:** Designed to support synchronous sessions, promote learning, and engage learners in real-time.
- **Case Studies:** Statements for exercises of real-life examples that show how algorithmic bias works, encouraging critical thinking and discussion.
- **Formative Assessment Exercises** require learners to articulate their understanding and apply concepts in writing to receive detailed feedback.
- **Expert Videos:** Curated videos offer industry leaders and academics insights into the topic area, giving the learner a wider understanding of the subject matter.
- **Project-Based Exercises:** Project exercises mimic actual applications of course materials, promoting practical skills and teamwork.

The course itself is built around key elements tailored specifically for this target audience:

- **Interactive eLearning Screens:** These screens enhance engagement by allowing learners to interact directly with the content, fostering a deeper understanding through active participation.
- **Interactive Flipbooks:** Flipbooks present information in a visually appealing and easily navigable format, incorporating multimedia elements to accommodate varied learning preferences.
- **Roulette Game with Multiple Choice Questions:** This digital game makes tests more exciting and, at the same time, involves learners enjoyably and interactively. The roulette game provides immediate feedback to students and helps them determine their level, making learning fun and effective.

The materials are tailored for instructional objectives and accessible on various devices. Prototyping and pilot testing involve internal checks and sample group feedback. An instructor training guide ensures efficient delivery and support systems aid students with technical or content issues. Synchronous sessions foster real-time interaction and cooperation. The "Algorithmic Bias" class, featuring a digital roulette game and interactive components, offers a comprehensive blended learning experience. A toolkit for trainers ensures effective delivery and continuous learner support.

3.2 Game-Based Learning

The CHARLIE project uses serious games as an innovative training strategy to educate and engage students on complex topics. These games increase motivation and facilitate learning by making it dynamic and interactive, especially for disadvantaged youth in STEM. They provide a safe environment for exploring and learning from mistakes, which is essential for developing critical and ethical skills. By connecting theory to practice, serious games demonstrate the relevance of ethical principles and technical skills, ensuring an inclusive, engaging learning environment to address AI and ethics education challenges.

Serious games enhance engagement and motivation by making learning dynamic and interactive, especially for disadvantaged youth in STEM. These games provide a safe space to explore and learn from mistakes, which is crucial for developing critical and ethical skills. They connect theory to practice, demonstrating the real-world relevance of ethical principles and technical skills. This ensures an inclusive and engaging environment, effectively addressing AI and ethics education challenges. The CHARLIE project fosters an innovative and effective pedagogical approach to complex topics by integrating serious games.

4. Results and Discussion

The developed the "Algorithmic Bias" course using a blended learning approach, which includes synchronous sessions, asynchronous eLearning, and a digital game for formative assessment, which will be pilot tested in three HE institutions, UIB, AU, and VAMK, in Spain, Denmark, and Finland, respectively, and will involve 90 participants. The course features interactive screens, flipbooks, quizzes, project-based exercises, expert videos, concept maps, and a digital roulette game to provide a comprehensive learning experience. We expect feedback to highlight the effectiveness of the interactive elements in enhancing understanding and retention. The digital roulette game aims to make learning enjoyable and provide immediate feedback, helping students gauge their comprehension in real-time.

4.1 Anticipated Results and Effectiveness of Game-Based Learning

The CHARLIE project anticipates significant benefits from its digital serious game, designed as an escape room with four themed rooms for youth aged 12-18. This game integrates EntreComp, DigComp 2.0, and GrenComp frameworks to introduce AI concepts, highlight biases, and explore ethical considerations. High engagement and positive reception are expected.

The scenario-based approach is anticipated to facilitate critical thinking and ethical reasoning, helping students understand algorithmic biases in everyday technology. Game-based learning will enhance ethical AI education by engaging students, fostering critical thinking, and promoting practical understanding. Immediate feedback and real-world applications help students internalize ethical principles and understand algorithmic biases.

To evaluate the effectiveness of our game-based learning approach, we will use a combination of quantitative and qualitative metrics. These performance metrics include pre- and post-assessment scores to measure knowledge gains, engagement levels to track student participation and interest, and participant feedback to gather insights into the learning activities' user experience and perceived value. Additionally, we will analyse the completion rates of the game-based modules and the time spent on each activity to understand learner engagement better. Surveys and focus groups will provide qualitative data on the learners' perceptions of the

effectiveness and enjoyment of the educational games. These comprehensive metrics will help us assess the overall success of the proposed setup and identify areas for improvement, ensuring that the game-based learning approach meets its educational objectives and effectively enhances students' understanding of ethical AI.

4.2 Broader Implications for Educational Practices

The CHARLIE project's innovative approach has broader implications for integrating game-based learning into ethical tech education. Digital games, when complemented by well-structured eLearning courses and comprehensive support for educators, could serve as powerful tools for teaching ethical concepts. The project's focus on creating linkages between HE, AE, and Youth sectors is crucial for the widespread adoption and recognition of ethical AI curricula. The interdisciplinary and collaborative approach facilitated by the project is expected to serve as a model for other educational initiatives aiming to address the ethical implications of emerging technologies.

While the CHARLIE project focuses on a European audience, the framework can be adapted for non-European environments by considering local cultural norms, educational systems, and technological access. For instance, additional support and training may be necessary to ensure effective implementation in regions with lower digital literacy. Moreover, cultural differences may affect how ethical issues are perceived and addressed, necessitating tailored content that resonates with local values and norms.

For example, in many Asian countries, collective societal values may influence how ethical considerations are taught and understood, contrasting with the individualistic approach often found in Western education. Thus, educators need to incorporate community-focused ethical scenarios to make the learning more relevant and impactful.

4.3 Challenges and Solutions

The CHARLIE project faced several challenges during its development and implementation phases. Addressing these challenges was crucial for success. Key challenges and solutions are summarized in Figure 4:

Figure 4: CHARLIE's Challenges and Solutions

Additionally, the technological infrastructure in non-European regions can vary significantly, impacting the delivery and accessibility of digital game-based learning tools. In areas with limited internet access or older technology, offline or low-bandwidth versions of educational games may be necessary. Collaborating with local educators to understand and address these challenges can help ensure the framework remains effective and inclusive. Adapting the CHARLIE project's framework for non-European environments involves a multifaceted approach considering cultural sensitivities, educational practices, and technological limitations. By doing so, the framework can be universally applicable, promoting ethical AI education globally.

Potential confounding factors that could impact the study's results include varying levels of digital literacy among participants, differences in educational infrastructure, and socio-economic disparities. These factors could skew the effectiveness of the game-based learning approach. To mitigate these issues, we will implement standardized training sessions to level the playing field, ensure that all participants have access to the necessary resources, and use control-group studies to isolate the effects of our intervention. By addressing these

confounding factors proactively, we aim to ensure the validity and reliability of our findings, providing a clear picture of the framework's effectiveness.

5. Implications and Future Work

Due to rapid developments in AI technology and the evolving ethical principles associated with them, it is crucial to update the knowledge and contents of the course continuously. While the game remains relevant, one must stay prepared for rapid advancements in this field. Continually updating the content and the game is essential to keep pace with these changes. Furthermore, games are advancing rapidly, as are the methods by which people engage with them, necessitating agile approaches to educational game design and implementation.

5.1 Educational Impact

In the project's current phase, before pilot testing the game, the expectations are primarily theoretical and aim to fulfil the predetermined learning outcomes. Numerous studies on the educational impact of game-based learning highlight its varied effects on students' motivation and ability to absorb information better. For instance, a study by Krath et al. (2021) examined 118 theoretical foundations and revealed that no single theory explains the effectiveness of game-based approaches. However, it did show that game-based learning has a high potential for improving effect and motivation, behaviour, and learning outcomes.

In the context of ethical AI and algorithmic biases, transforming ethical principles, frameworks, and guidelines into practical tools is challenging, and many frameworks lack contextual application (Morley et al., 2020; Prem, 2023). Morley et al.'s (2020) study revealed a gap in tools that effectively address the ethical impact of AI on individuals or society, possibly due to the complexity of translating human behaviours into universal tools. Introducing a gaming concept can bridge the gap between abstract ethical principles and practical application, helping students grasp the real-world implications of these issues on individuals and society.

The effectiveness of the proposed method may vary across different demographic groups, such as age, social status, gender, and prior exposure to technology. Learning styles, cultural backgrounds, and educational experiences can significantly influence learning outcomes. To address this, we plan to conduct targeted studies that examine these variations in detail. By understanding diverse groups' specific needs and challenges, we can tailor our educational approach to ensure inclusivity and equity. This may involve developing different versions of the educational games, providing additional support for certain groups, and creating culturally relevant content to enhance engagement and effectiveness across all demographics.

5.2 Implications for Future AI and Tech Education

A primary goal is to incorporate a human-centered approach into teaching technical subjects. Project Charlie aims to enhance empathy-based methods in AI and technology education and boost proactive efforts in fostering individuals' sense of responsibility and civic engagement. Utilizing a game-based approach allows for engaging a broader audience within higher education and across various age and educational groups. Introducing ethical considerations through a human-centered approach in future AI and technology education represents a strategy to embed ethical principles into AI solution development from the initial stages.

6. Conclusion and Future Directions

The CHARLIE project's game-based approach primarily targets individuals aged 12 and up, but rapid digitalization exposes children to AI technologies at much younger ages. Introducing ethical considerations early could enhance understanding and foster inherent empathy and ethical sensibilities, proving beneficial in the long term, significantly if these individuals eventually develop AI solutions. Through serious gaming, users can test the consequences of their actions on others and the environment, thereby gaining empathy.

Despite the advantages of game-based learning, critical assessment is needed to determine its effectiveness in increasing empathy. While some studies suggest potential desensitization from violent game content (Andersson & Bushman, 2001), others indicate that game-based approaches can enhance empathy (Wulansari et al., 2020). Comparing empathy levels elicited by different media—movies, books, and games—could provide insights into creating engaging content. As education increasingly adopts digital solutions, sophisticated approaches are necessary to engage effectively.

Our study needs to account for the potential variances in adopting game-based learning to teach AI ethics at the country level. For example, disparities in digital skills among European children (Mascheroni et al., 2020) could lead to unequal opportunities for technological advancements. Addressing these disparities and cultural differences is essential when designing educational games to ensure effective connectivity and empathy towards avatars and game elements.

The CHARLIE project's game-based efforts are directed toward young women and youths from socioeconomically disadvantaged backgrounds. While engaging these groups is crucial, it is equally essential to ensure that those leading AI development—often not from disadvantaged groups—consider the ethical implications for minorities disproportionately affected by these technologies. Reaching this target group could significantly enhance the project's impact. By fostering early ethical education and inclusivity, the CHARLIE project can ensure that future AI developments are empathetic and socially responsible, paving the way for a more equitable technological landscape.

References

Abulibdeh, A., Zaidan, E., & Abulibdeh, R. (2024). Navigating the confluence of artificial intelligence and education for sustainable development in the era of industry 4.0: Challenges, opportunities, and ethical dimensions. Journal of Cleaner Production, 140527. https://doi.org/10.1016/j.jclepro.2023.140527

Anderson, C, Bushman, B. (2001) Effects of Violent Video Games on Aggressive Behavior, Aggressive Cognition, Aggressive Affect, Physiological Arousal, and Prosocial Behavior: A Meta-Analytic Review of the Scientific Literature doi/10.1111/1467-9280.00366

European Commission. (2019.). Ethics guidelines for trustworthy AI. Digital Strategy. Retrieved [date], from https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai

Fedele, A., Punzi, C., & Tramacere, S. (2024). The ALTAI checklist as a tool to assess ethical and legal implications for a trustworthy AI development in education. Computer Law & Security Review, 53, 105986. https://doi.org/10.1016/j.clsr.2024.105986

Kayyali, M. (Ed.). (2024). Building Resiliency in Higher Education: Globalization, Digital Skills, and Student Wellness: Globalization, Digital Skills, and Student Wellness. IGI Global.

Khan, H. U., Malik, M. Z., Alomari, M. K. B., Khan, S., Al-Maadid, A. A. S., Hassan, M. K., & Khan, K. (2022). Transforming the capabilities of artificial intelligence in GCC financial sector: a systematic literature review. Wireless Communications and Mobile Computing, 2022. https://doi.org/10.1155/2022/8725767

Krath, J., Schürmann, L., Harald F.O. von Korflesch, H. (2021) Revealing the theoretical basis of gamification: A systematic review and analysis of theory in research on gamification, serious games and game-based learning, Computers in Human Behavior, 0747-5632 https://doi.org/10.1016/j.chb.2021.106963.

Lang, Josephine. "Workforce upskilling: can universities meet the challenges of lifelong learning?." The International Journal of Information and Learning Technology 40, no. 5 (2023): 388-400. DOI: 10.1108/IJILT-01-2023-0001

Mascheroni, G., Cino, D., Mikuška, J., Lacko, D., & Smahel, D. (2020). Digital skills, risks and wellbeing among European children: Report on (f)actors that explain online acquisition, cognitive, physical, psychological and social wellbeing, and the online resilience of children and young people. Zenodo. https://doi.org/10.5281/zenodo.4267977

Morley, J., Floridi, L., Kinsey, L., & Elhalal, A. (2020). From What to How: An Initial Review of Publicly Available AI Ethics Tools, Methods and Research to Translate Principles into Practices. Science and Engineering Ethics, 26(4), 2141–2168. https://doi.org/10.1007/s11948-019-00165-5

Nguyen, A., Ngo, H. N., Hong, Y., Dang, B., & Nguyen, B. P. T. (2023). Ethical principles for artificial intelligence in education. Education and Information Technologies, 28(4), 4221-4241. https://doi.org/10.1007/s10639-022-11316-w

Prem, E. (2023). From ethical AI frameworks to tools: A review of approaches. AI and Ethics, 3(3), 699–716. https://doi.org/10.1007/s43681-023-00258-9

Schwartz, R., Schwartz, R., Vassilev, A., Greene, K., Perine, L., Burt, A., & Hall, P. (2022). Towards a standard for identifying and managing bias in artificial intelligence (Vol. 3, p. 00). US Department of Commerce, National Institute of Standards and Technology. https://doi.org/10.6028/NIST.SP.1270

Slimi, Z., & Carballido, B. V. (2023). Navigating the Ethical Challenges of Artificial Intelligence in Higher Education: An Analysis of Seven Global AI Ethics Policies. TEM Journal, 12(2).

UNESCO. (2021). Recommendation on the ethics of artificial intelligence. Retrieved from https://unesdoc.unesco.org/ark:/48223/pf0000380455

Wulansari, O, Pirker, J, Kopf, J, Guetl, C, Video Games and Their Correlation to Empathy: How to Teach and Experience Empathic Emotion DOI:10.1007/978-3-030-40274-7_16