



NEM SUMMIT 2024

AI for Virtual World

Explainable AI in VR Gaming: Enhancing Transparency and Trust

Dr. Makrina Viola Kosti - CERTH



CERTH
CENTRE FOR
RESEARCH & TECHNOLOGY
HELLAS



**Information
Technologies
Institute**



This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101132449

NEM Summit, 23-24 October 2024, Brussels-Belgium



Introduction

- **Role of Explainable AI (XAI) in virtual reality gaming**
- **Challenges**
- **Problems**
- **Benefits for the gaming culture**
- **Benefits for VR ecosystem**
- **Policy implications for EU Commission stakeholders**





What is Explainable AI (XAI)?

Definition: Explainable AI refers to methods and techniques that make the output of AI systems clearer and more understandable to humans.

Purpose: To ensure transparency and trust in AI decisions by providing clear explanations of how predictions are made.



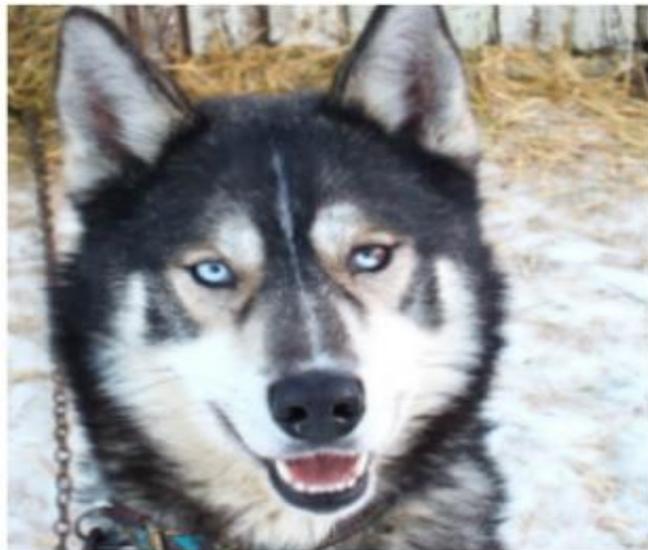
*Gunning, D. (2017). Explainable Artificial Intelligence (XAI). *Defense Advanced Research Projects Agency (DARPA)*



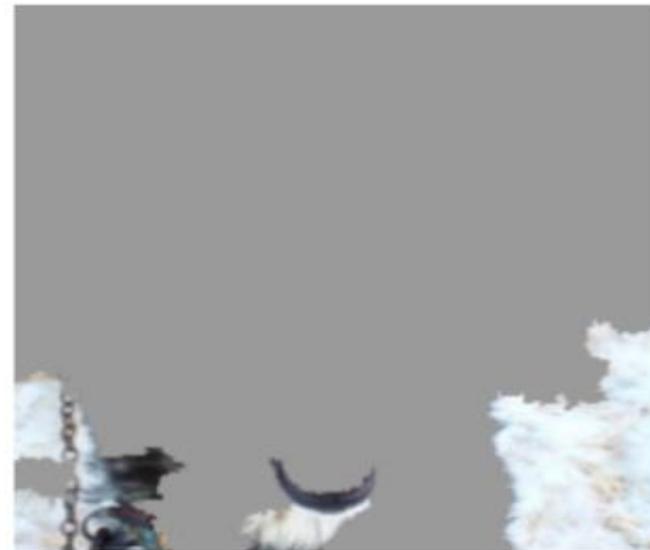


Are Huskies Wolves?

- **Wolf Images:** In the training dataset, most images of wolves had snowy backgrounds.
- **Husky Images:** Images of huskies typically did not feature snow.



(a) Husky classified as wolf



(b) Explanation





XAI Goals

XAI aims to solve the following problems:

- **Right to explanation:** In scenarios where AI systems act autonomously and have various impact, XAI ensures legal compliance with international regulations, such as AI Act.
- **User Trust:** XAI enhances user trust by providing clear and understandable explanations of AI decisions to non-experts, facilitating AI system adoption.
- **Insight Discovery:** XAI helps uncover meaningful relationships and correlations between input data and model outputs, leading to improved future solutions.
- **Error Analysis:** XAI helps in identifying biases or inconsistencies in data, building more robust and fair AI models.





XAI Goals

Questions related to AI systems

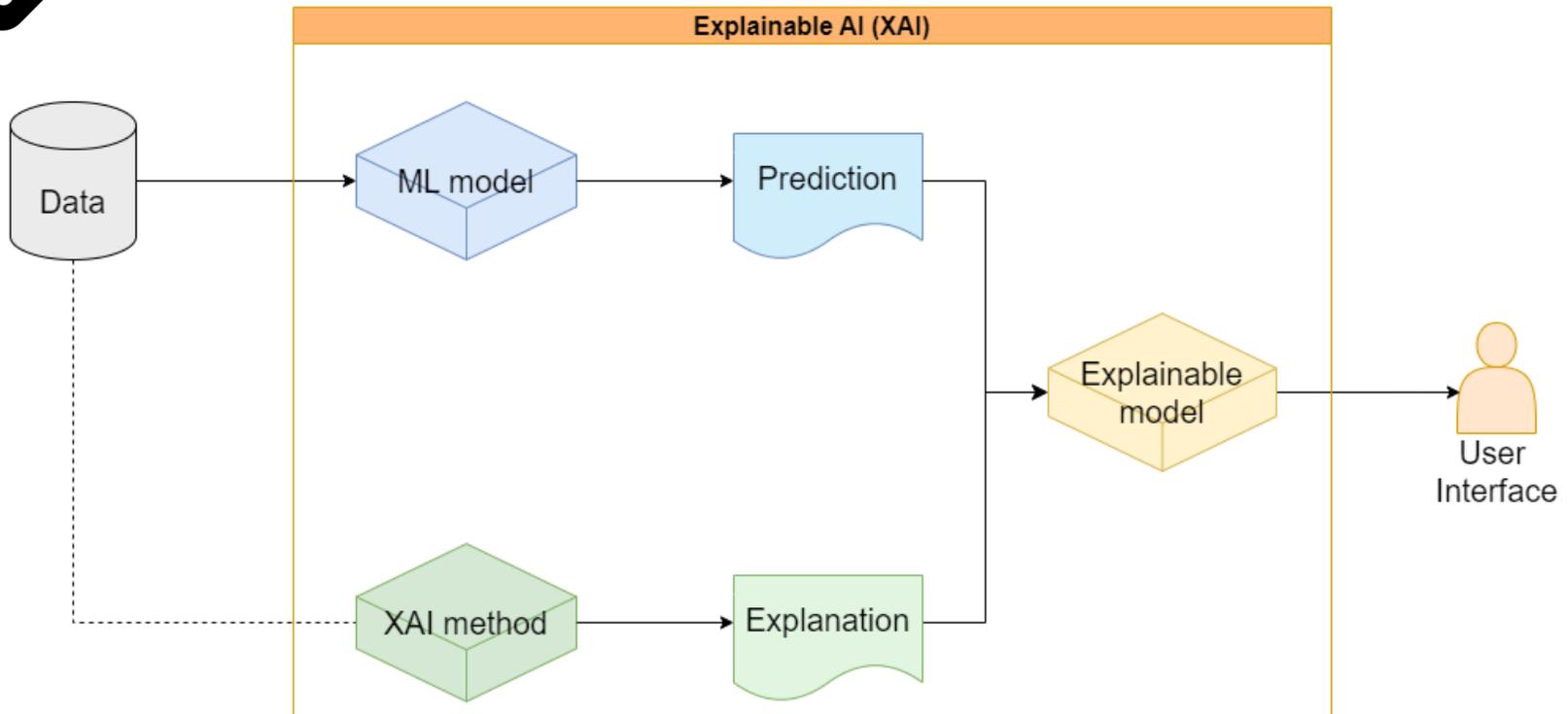
Explainable AI aims to answer the following questions:

- 1) **How** did the AI system reach its **decision**?
 - Which input features were most important?
 - What was the model's reasoning process?
 - What patterns has the system learned?
 - What training data influenced the decision?
- 2) **Why** was this decision **made** instead of a different one?
 - What alternative decisions were considered?
 - What factors led to this particular choice?
- 3) **What** is the **uncertainty** level of the system?
 - How confident is the model in its decision?
 - What is the probability of potential errors?





Explainable AI for Users



Common Objectives:

- Provide real-time, easy-to-understand explanations to users
- Enhance user understanding and trust of the AI model
- Potentially enable user interaction with the interface





Gameplay experience

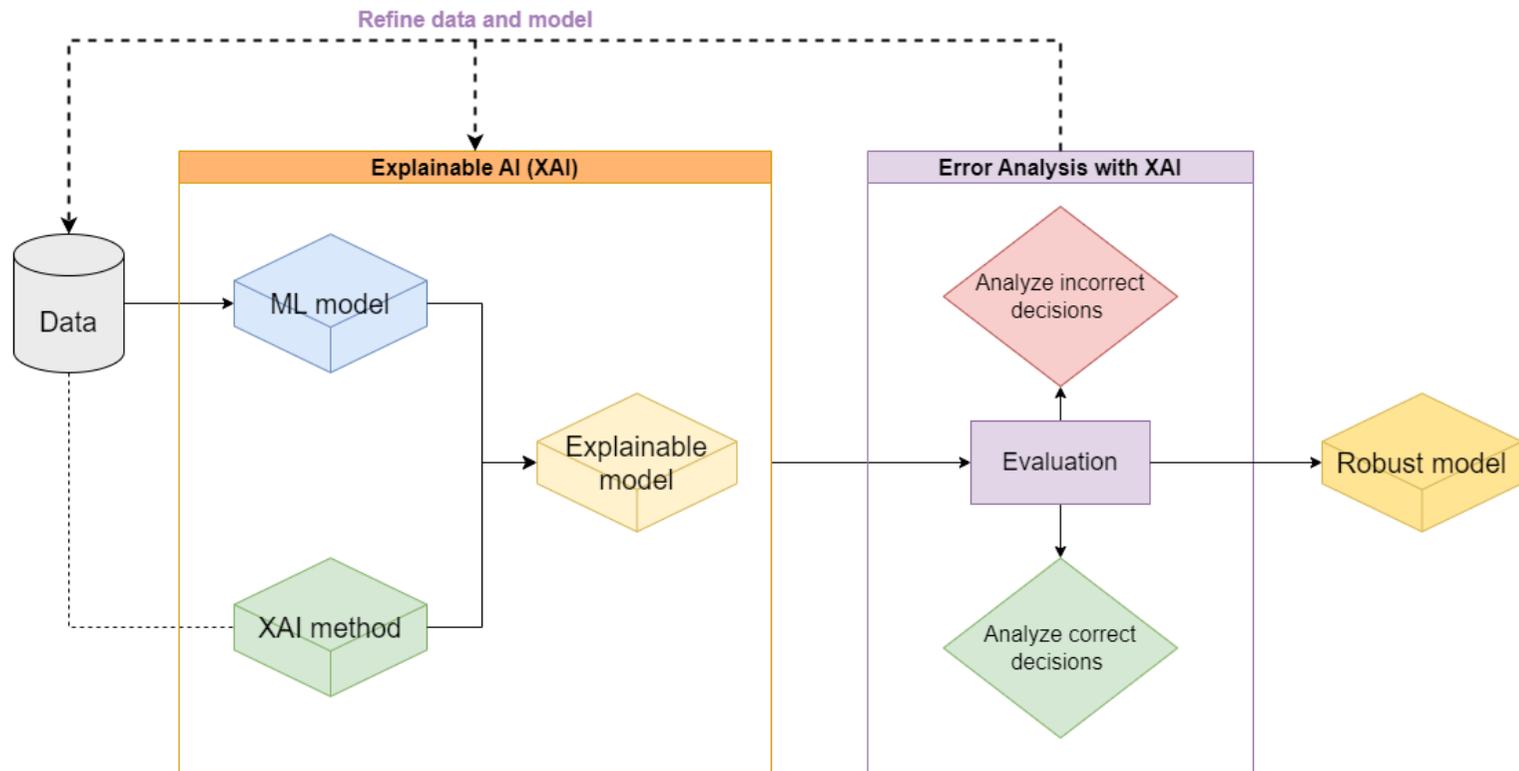
- **Personalized Experiences:** AI can adapt to player preferences more effectively when its decision-making process is transparent.
- **Adaptive Learning:** By explaining AI adjustments, players can better understand and improve their own gameplay strategies.
 - Imagine a VR game where the AI explains why it adjusted the difficulty level, helping players feel more in control and engaged.

* Lopes, P., Catarino, A., & Prendinger, H. (2019). The Effect of Explainable AI on Player Trust and Satisfaction in a Game Setting. *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*, 1-12.





Explainable AI for Developers



Common Objectives:

- Provide clear explanations to the developers
- Build a more robust and trustworthy AI system for end users

Error analysis with XAI:

- Analyze correct decisions to verify the model's reasoning
- Examine incorrect decisions to identify and address model and data faults





Facilitating Developer Insights

- **Debugging and Development:** Transparent AI systems make it easier to identify flaws and unintended behaviours during development.
- **Innovative Design:** Understanding AI behaviours can inspire new gameplay mechanics and features.
 - This collaborative understanding between AI specialists and game designers can lead to more innovative and engaging games.

* Elbert, Nico, and Christoph M. Flath. "Process Mining for Game Analytics." *2024 IEEE Conference on Games (CoG)*. IEEE, 2024.





Technical challenges

- **Integration Complexity:** Adding explainability features can strain computational resources. VR games require high frame rates to maintain immersion, and additional computations for XAI can cause performance issues.
- **Performance Constraints:** Developers must optimize XAI algorithms to function in real-time without causing lag or latency.
- **Balancing Immersion and Explanation:** Providing explanations should not disrupt the player's immersion. We need intuitive ways to present information without pulling players out of the experience.
 - Use simplified models

*Ali, Sajid, et al. "Explainable Artificial Intelligence (XAI): What we know and what is left to attain Trustworthy Artificial Intelligence." *Information fusion* 99 (2023): 101805.





Ethical and Social Challenges

- **Bias and Fairness:** AI models trained on biased data can exhibit discriminatory behaviours. In gaming, this could mean NPCs reacting differently to players based on gender, ethnicity, or other characteristics, which is unacceptable.
- **Inclusivity:** Ensuring that AI interactions are respectful and appropriate for all users is essential for creating an inclusive gaming environment.
- **User Privacy:** Personalized AI experiences require data collection. We must handle personal data responsibly to protect user privacy.
- Addressing these challenges requires diverse datasets and continuous monitoring to mitigate biases.

* Noble, S. U. (2018). *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York University Press.





Regulatory Challenges

- **Data Protection Laws:** Compliance with regulations like the General Data Protection Regulation (GDPR) is mandatory. GDPR emphasizes user consent and data minimization.
- **EU AI Act:** The proposed EU AI Act aims to regulate AI systems, focusing on transparency, safety, and accountability.
- **Certification and Standards:** Meeting industry standards ensures that AI systems are reliable and trustworthy.
 - Developers must proactively engage with these regulations to avoid compliance issues and contribute to shaping reasonable policies.

*European Commission. (2021). Proposal for a Regulation laying down harmonised rules on artificial intelligence (Artificial Intelligence Act).





Regulatory Examples

Personalized AI Explanations:

- **Scenario:** An AI explains why a player's in-game character reacted a certain way based on past choices.
- **Compliance:** The game collects only the choices relevant to that explanation and informs the player about this data collection.

Behavioural Analytics:

- **Scenario:** The AI adjusts game difficulty by analyzing player performance metrics.
- **Compliance:** Players are informed about this data usage and consent to it; data unrelated to performance is not collected.





Gaming Culture - Impact

- **Representation:** AI that respects and reflects diverse cultures enhances player connection and satisfaction.
- **Global Reach:** Culturally sensitive AI can make games more accessible and appealing to a worldwide audience.

*Mukherjee, S. (2018). Video Games and the Future of Cultural Representation. *Journal of Comparative Research in Anthropology and Sociology*, 9(1), 1-12.





Advancement of the VR Ecosystem

- **Innovation Catalyst:** It encourages the development of new technologies and methodologies within the industry.
- **Industry Leadership:** Early adopters of XAI set industry standards and can lead the market.
 - Moreover, advancements in XAI for gaming can have cross-industry applications, benefiting fields like healthcare, education, and beyond.





Societal Impact

- **Educational Applications:** VR combined with XAI can be used for training simulations, providing clear insights and explanations.
- **Ethical AI Practices:** Implementing XAI sets a precedent for responsible AI use in consumer products, promoting ethical standards.
 - For instance, VR training programs for surgeons can use XAI to explain procedural outcomes, enhancing learning experiences.



Image: https://www.medica-tradefair.com/en/digital-health/Surgery_in_3D_Virtual_Reality_in_the_OR





Policy Implications

- **Need for Supportive Policies:** Policies that encourage innovation while ensuring ethical standards will facilitate XAI integration.
- **Risk Mitigation:** Regulations can help prevent misuse of AI technologies, protecting consumers and the industry.
 - Policymakers have a crucial role in fostering an environment where XAI can thrive, balancing innovation with necessary safeguards





Conclusion

Explainable AI addresses numerous challenges in VR gaming
from **technical**
ethical issues
to **societal** impacts.

It enhances user **trust**,
improves **gameplay** experiences,
and fosters **innovation** within the VR ecosystem.





Explainable AI in VR Gaming: Enhancing Transparency and Trust

Thank You!



CERTH
CENTRE FOR
RESEARCH & TECHNOLOGY
HELLAS



⇒ Follow us on social media



This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101132449

NEM Summit, 23-24 October 2024, Brussels-Belgium