Developing Trustworthy Educational Metaverses Using Open-Source Generative AI (GenAI) and Mixed Reality (MR) Technologies

Prof. Junaid Qadir[1], Prof. Ala Al-Fuqaha[2], Prof. Muhammad Bilal[3], Dr. Sajjad Hussain[4]

[1]Qatar University, Qatar

[2]Hamad bin Khalifa University, Qatar

[3]Birmingham City University, United Kingdom

[4] University of Glasgow, United Kingdom

IEEE EDUCON 2024

IEEE Global Engineering Education Conference

Presentation Outline

1. Introduction to Metaverse
2. Introduction to Educational Metaverses

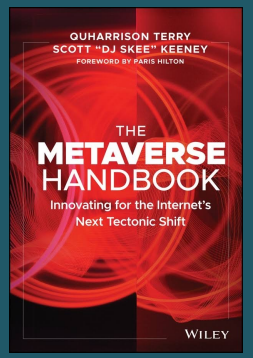
- *Education in the Metaverse: Opportunities and Risks*
- *An Introduction to Generative AI (GenAI) and Open-Source GenAI*
- *An Introduction to Mixed Reality (MR) and Open-Source MR*
- *An example case study incorporating GenAI/ Metaverse.*

1

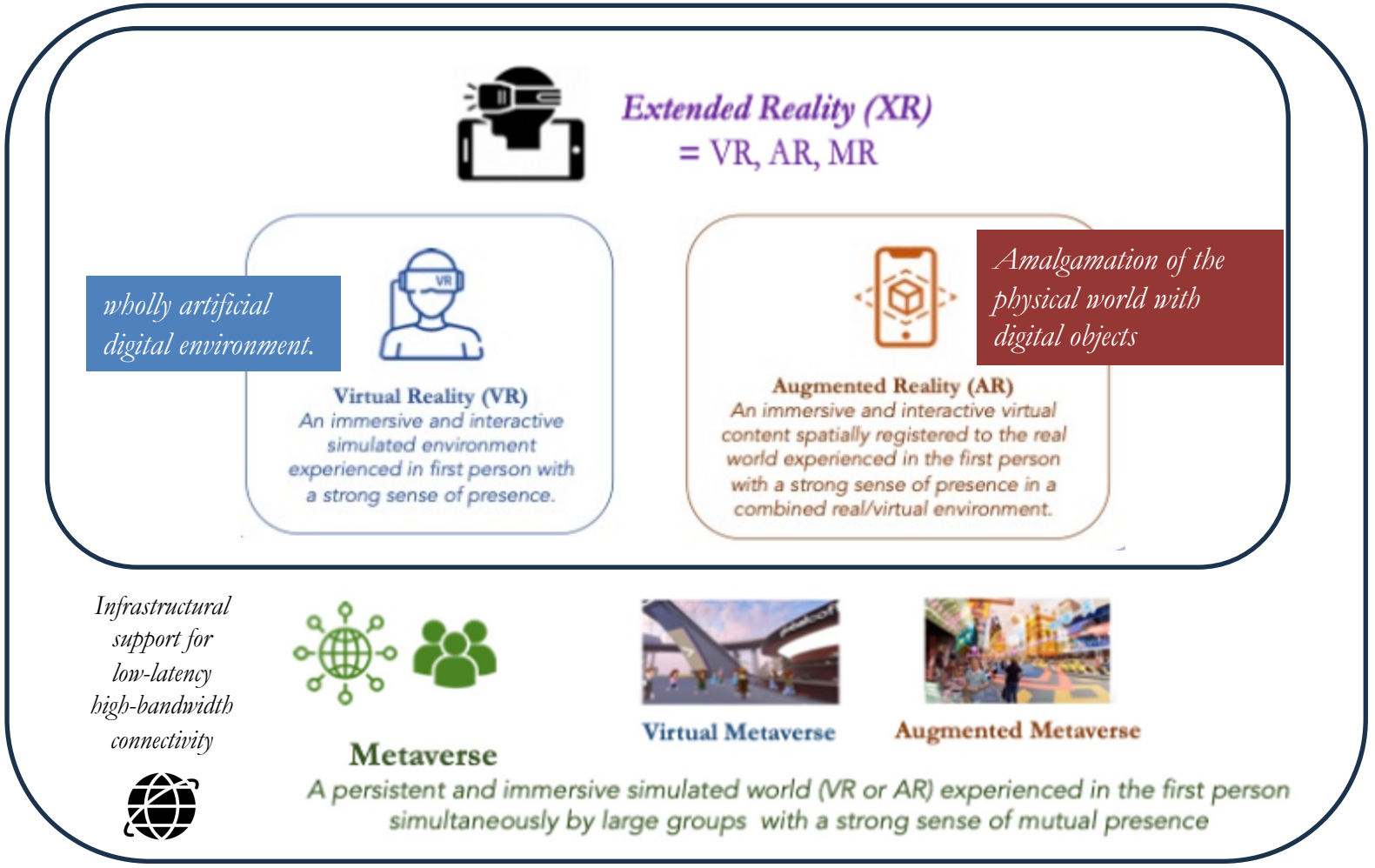
Introduction

Defining the Metaverse

The Metaverse is a culmination of many applications of technology. There is no one single thing that makes the Metaverse the Metaverse



Introduction to AR/VR and the Metaverse



Enabling Technologies:

- *Spatial Computing*
- *Spatial Web*
- *Haptic Technology*
- *Localization and Mapping*
- *AR/VR Libraries*
- *Gaming Industry*

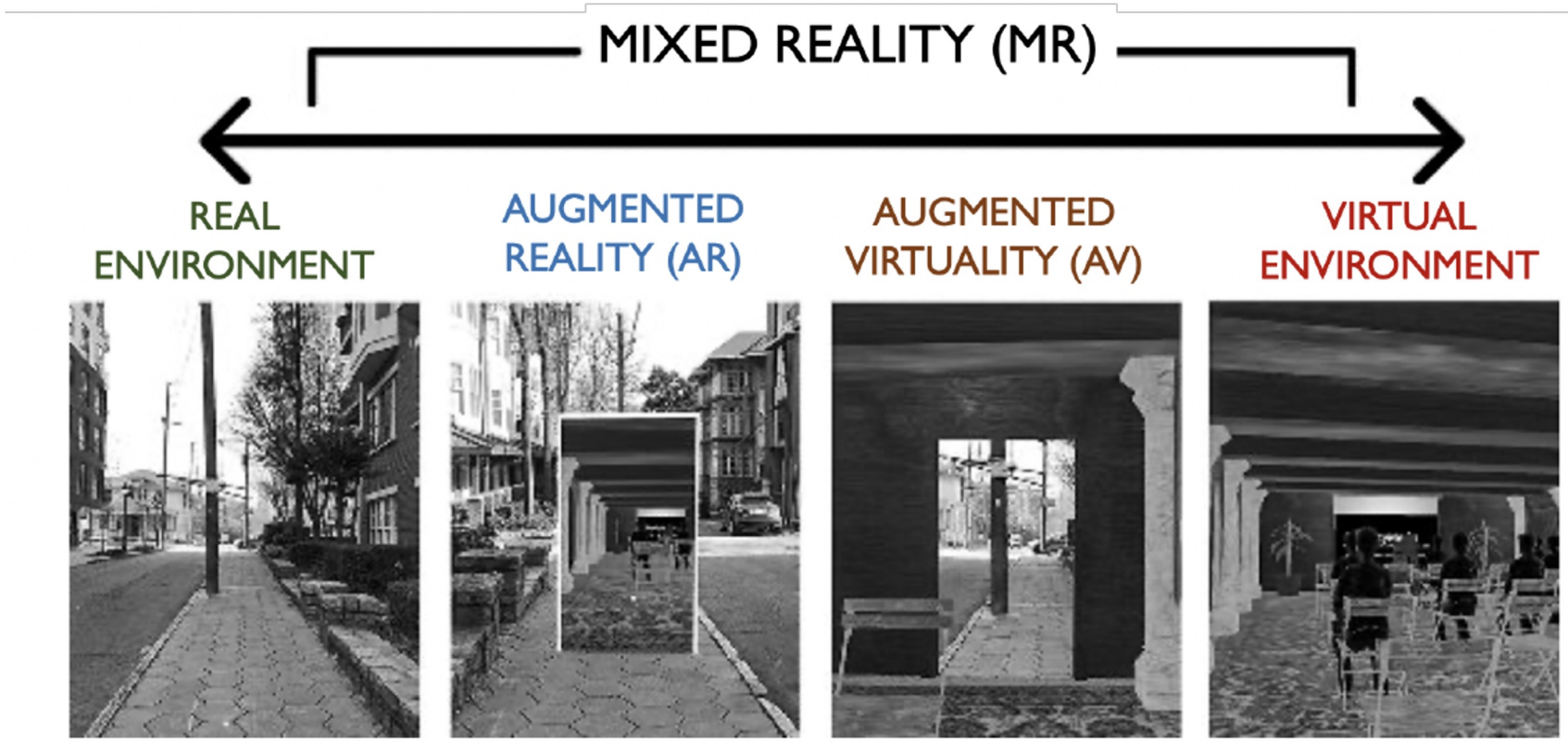
Meta

Vision Pro

Mixed Reality (MR) Definition

The term MR describes a spectrum encompassing both Augmented Reality (AR) and Virtual Reality (VR). In MR environments, physical and digital objects coexist and interact in real time. Users can engage with virtual objects while still being aware of and interacting with the real world around them.

Reality to Virtuality Continuum



Spectrum ranging from Reality to Virtuality [Milgram and Kishino (1994) continuum.]
Figure adapted from "Reality Media" (Source: Bolter, Engberg, & MacIntyre, 2021)

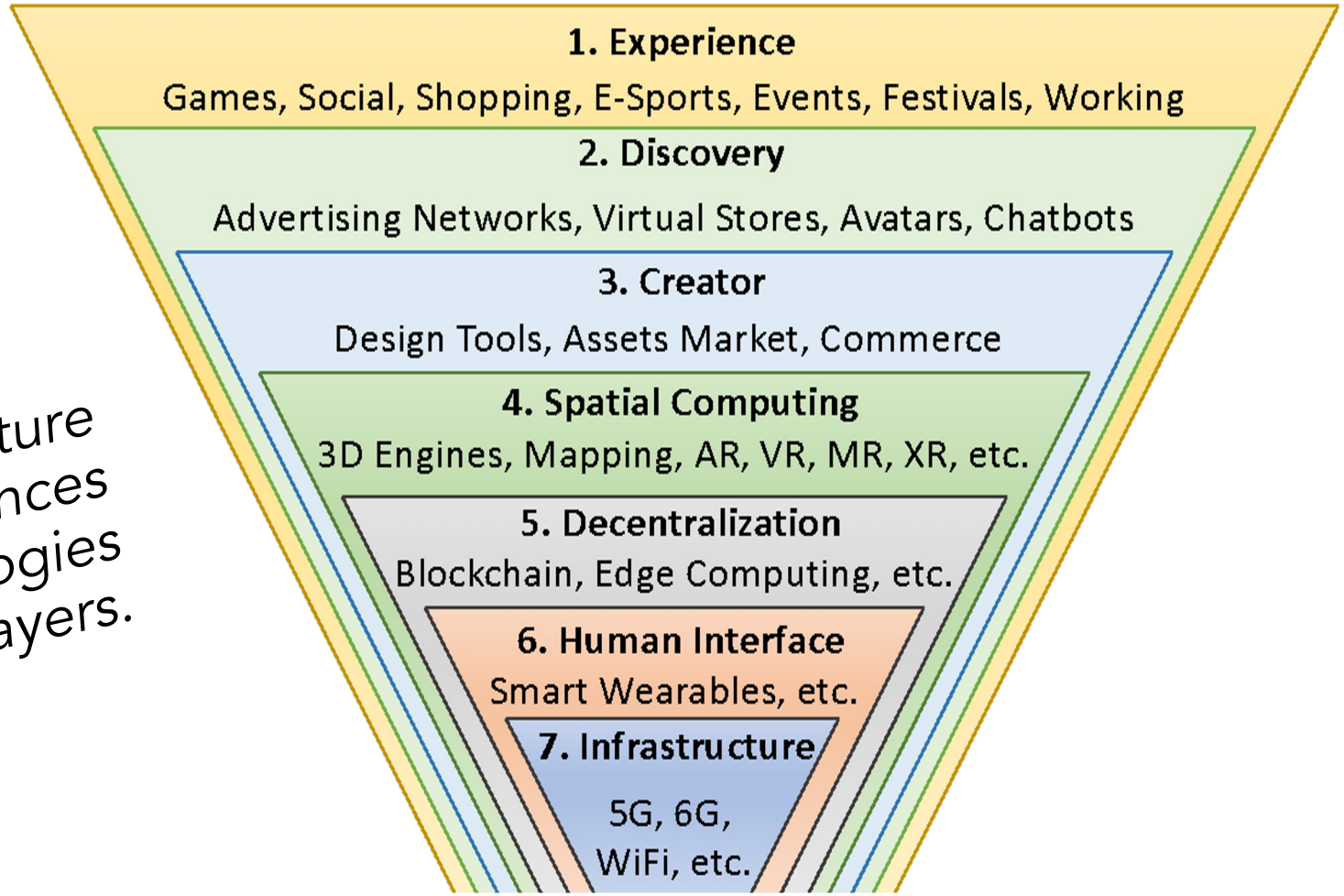
Metaverse Enabling Technologies

Metaverse is the next generation of the Internet that will surround us both graphically and socially enabled by various immersive and virtual technologies including XR.

Qayyum, A., Butt, M. A., Ali, H., Usman, M., Halabi, O., Al-Fuqaha, A., ... & Qadir, J. (2024). Secure and trustworthy artificial intelligence-extended reality (AI-XR) for metaverses. ACM Computing Surveys, 56(7), 1-38.

Architectural Overview of Metaverse

Metaverse's architecture expands from people's experiences to underlying enabling technologies and has seven layers.



Qayyum, A., Butt, M. A., Ali, H., Usman, M., Halabi, O., Al-Fuqaha, A., ... & Qadir, J. (2024).
Secure and trustworthy artificial intelligence-extended reality (AI-XR) for metaverses. ACM Computing Surveys, 56(7), 1-38.

Applications of Metaverses

Applications of Metaverses

Industry

- Manufacturing (i.e., developing better products and fault identification)
- Employees and Workers Training to Improve Safety
- Improving Planning for Manufacturing (Plants and Equipment)

Military

- Tactical Augmented Reality
- Immersive (Synthetic) Training Environments
- Creating physically and psychologically intensive combat situations, ...

Ecommerce

- Virtual Markets and Business
- Virtual Assists
- Virtual Shopping
- AR and VR assisted immersive shopping

Education

- Improving teaching using immersive environment
- Immersive learning environments for students
- Personalized learning interventions, ...

Healthcare

- Planning and Training
- AR-assisted Surgical Interventions
- Pharmaceuticals
- Vein Visualization, ...

Real-Estate

- Simulating 3D architectures
- Realistic and Immersive Virtual Tours
- Customized Tours (as per customer needs)
- Adding narrations alongside ambient music, light, & sound effects, ...

Gaming & Entertainment

- Immersive Gaming
- Enhanced Entertainment and Social Networking
- Virtual Meetings and Workspaces, ...

How educational metaverses can help improve education?

Accessible Learning

Foster accessible learning through inclusive, immersive environments, catering to diverse styles and ensuring equitable access to resources.

Remote Learning

Immersive online environments, allowing students to collaborate, engage, and access educational content remotely.

Personalized Learning

Personalized learning through tailored content and experiences to individual needs, fostering adaptive and customized educational journeys.

Coummuniations

improved communication through interactive platforms fostering real-time engagement and effective information exchange among students and educators.

Immersive Learning

Learners can access personalized and adaptive content, fostering a gamified and interactive approach to education

Interactive Visualizations

Interactive visualizations for dynamic, immersive learning, enhancing comprehension through engaging and visually rich content.

Collaborative Learning

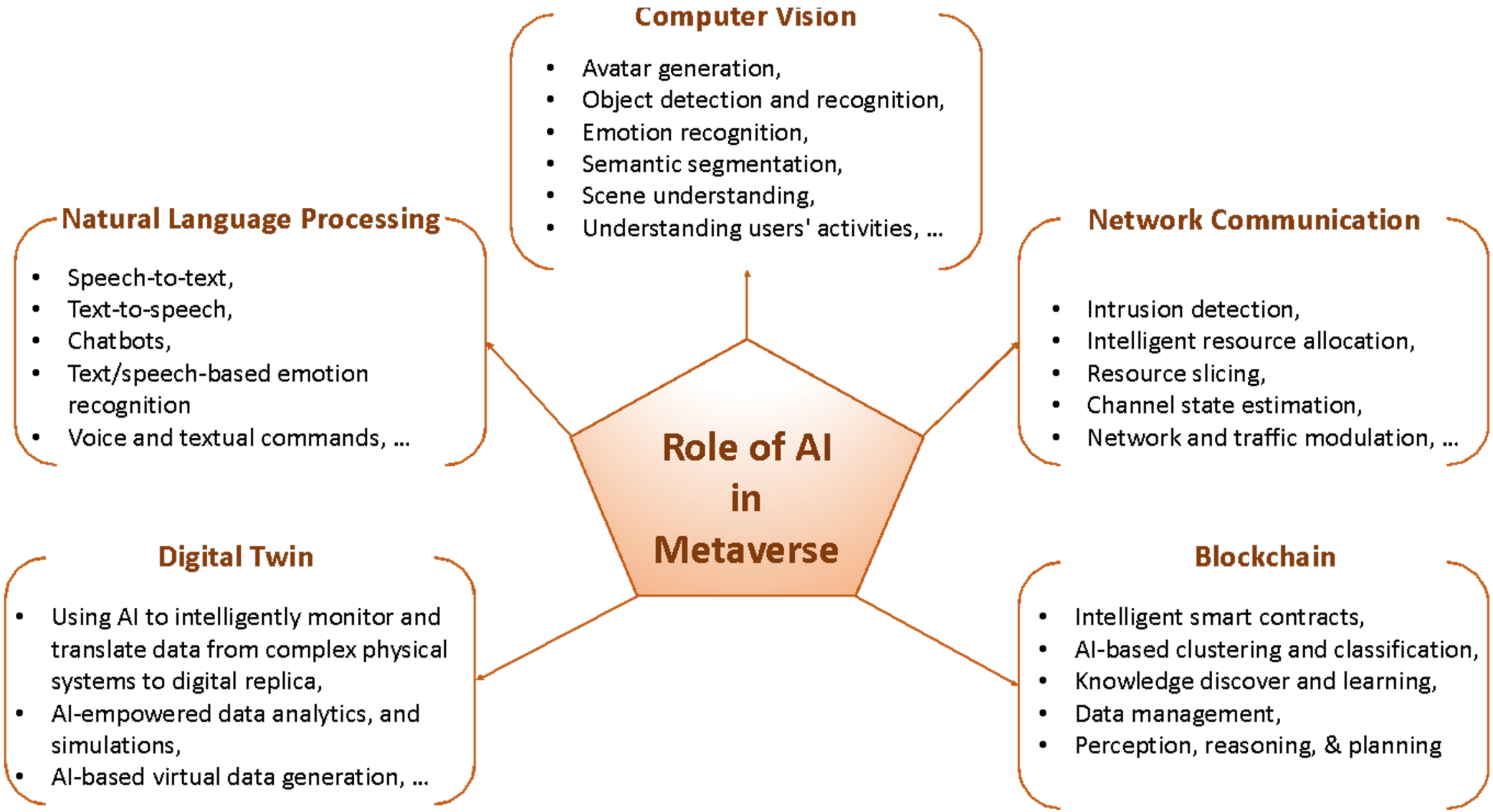
Empower collaborative learning by providing immersive virtual spaces where students globally can interact, share knowledge, and engage in dynamic, real-time educational experiences

Assessment and Analysis

Provides dynamic tools for evaluating student progress, understanding learning patterns, and tailoring personalized feedback for effective educational outcomes.

Abdullakutty, F., Qayyum, A., & Qadir, J. (2024). Trustworthy AI for Educational Metaverses. Authorea Preprints.

Artificial Intelligence's Role in Metaverses



2

An Introduction to Educational Metaverses

Education in the Metaverse

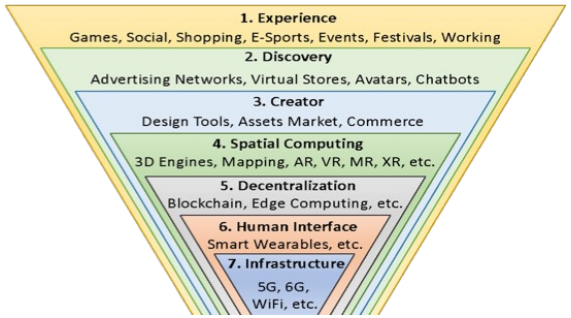
Future of Education: Immersive Teleportation Systems

Immersive teleportation systems redefine education by transcending physical limitations and fostering global connectivity.

- **Availability:** Live surgical training becomes globally accessible.
- **Accessibility:** Students worldwide join classes anytime, anywhere.
- **Communication:** Virtual classrooms enable seamless collaboration across borders.

Educational Metaverse: Opportunities and Risks

Security Issues in the AI-XR Metaverses



Different Layers of Metaverse

	Infrastructure	Human Interface	Decentralization	Spatial Computing	Creator Economy	Discovery	Experience
Enabling Technologies	- 5G, 6G - WiFi - Cloud - Data Center/Clusters, ...	- Smart Phones - Smart Watches - Smart Glasses - IoTs Devices - Wearables Devices, ...	- Edge Computing - AI agents - Blockchain - Microservices - Distributed Computing - NFTs, ...	- 3D Engines - VR, AR, MR, XR - Geospatial Mapping - Voice, Gesture, & Object Recog. - Data Integration, ...	- Design Tools - Assets Market - E-commerce - Workflow, ...	- Advertising Networks - Virtual Stores - Social Curation - Ratings - Avatars, Chatbots - Community-driven content, ...	- Games, E-Sports - Social Networks - Shopping - Festivals, Events - Learning - Working - Real time presence, ...
ML Use Cases/Applications	- Resource Allocation - Load Balancing - Spectrum Monitoring - Channel Estimation - Traffic off-loading - Attack Prevention - Network Fault Detection - Scheduling, ...	- Data Storage and Analytics - Mobile Phone-based sensing - Voice Recognition - Human-computer Interaction - Brain Computer Interfacing - Face and Emotion Recognition, ...	- Distributed ML - Federated Learning - Embedded ML - ML-based Authentication - ML-based Smart Contracts - Swarm Learning, ...	- 3D Reconstruction - Voice & Object Recognition and Tracking - Image Restoration - Image Super-resolution - Pose Estimation & Action Recognition - Object Detection & Segmentation, ...	- ML-based Graphics Generation - Simulating Real Objects - 3D Objects Rendering - Denoising - ML-based Virtual Characters and Avatar Generation, ...	- Smart Advertising - Generative Models - Conversational AI - Natural Language Processing - ML-based Personalization - Smart Video Streaming and Quality Assessment, ...	- Computer Vision - 3D Modeling and Mapping - Voice, Face, Emotion, & Gesture Recognition - Object Detection & Tracking - Analyzing Body Movements & Physical Interactions, ...
ML Associated Security Issues	**Adversarial ML Attacks on** - Network Traffic Classification - Intrusion Detection - Resource Allocation - SPAM Detection - Parameters Selection - Modulation Classification, ...	**Adversarial Attacks on ML-based** - Recognition Systems & Data Processing Pipeline - Data Poisoning and Manipulation - Privacy Violations and Breaches - Fake Commands, ...	**Adversarial ML Attacks on** - Model being trained in Federated and Distributed Learning - Attacks on Embedded ML, ...	**Physical World Attacks on** - Object and Voice Recognition - Action Recognition - Attacks on Semantic Segmentation - Object Detection and Tracking, ...	**Privacy and Adversarial ML Attacks** - User Profiling - Privacy Breaches - Identity Theft - Authentication Attacks on ML-based systems - Verification Systems, ...	**Adversarial ML Attacks on** - Attacks on Natural Language Processing - Fake Commands Generation - Deep Fakes Generation, ..	**Physical World Attacks on** - Object ,Voice, Face, Action, and Emotion Recognition - Attacks on Semantic Segmentation - Object Detection and Tracking, ...

Common ML Security Issues: (1) Evasion Attacks; (2) Adversarial ML Attacks; (3) Data and Model Privacy; (4) Reconstruction Attacks; (5) Model Poisoning or Stealing; (6) Extraction and Exploration Attacks; (7) Trojan Attacks; (8) Model Reuse Attack; (9) Membership Inference Attack ;(10) Inversion Attacks;

Common Attack Surfaces: (1) Manipulating Data Collection and Preprocessing Pipelines; (2) Tampering Trained/Deployed Models or their Training; (3) Modifying Model Outputs; (4) Attacks on Authentication Systems; (4) Distribution Shifts/Drifts;

Towards Trustworthy AI in the Educational Metaverse

Secure and Private AI

- Differential Privacy
- Federated Learning
- Cryptographic Techniques
- Adversarial ML Defenses

Controlled Text Generation

- Retrieval Augmented Language Modelling
- Pre-trained Text Representations-based methods
- Attribute alignment

Explainable AI

- Creating Intrinsic Models
- Creating Post-hoc Models
- Model-specific Explanations
- Model-agnostic Explanations
- Surrogate Methods
- Visualization Methods (e.g., GradCAM and IntegratedGrad).

Ethical AI

- Following Ethical Guidelines
- Human Centric AI Development
- Using Fair (bias free) Training Data
- Ensuring Security and Robustness
- Preserving Data & Model Privacy
- Enhancing Tranperency in AI predictions

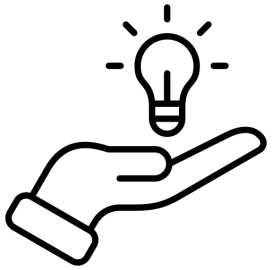
AR/VR Tech for Education

https://www.civicspace.tech/technologies/augmented-reality-virtual-reality/

Opportunities with AR/VR Tech for Education

Risks with AR/VR Tech for Education

Governance of AR/VR Tech



Opportunities

Education

Unique first-person immersive experience to enhance human perception and educate in the relevant environment.

AR/VR can be used for emotional support, and for managing stress and anxiety

Emotional Support

Accessibility and Inclusiveness

Customizaion for visually-impaired people with enhanced emphasis on other senses with support from haptic or audio interface.

Empathy and Awareness

Increase empathy and promote awareness of problems such as racism, war, and discrimination

Case Studies

SDG Action Campaign

VR For Democratizing Access to Education

Spotlight on Refugees/ War Victims

Participatory Governance

Awareness of Sexual Harassment

Awareness of Discrimination

Maternal Health Training

Awareness of Dire Poverty

Risks

Lack of Inclusiveness & Inequality Amplification

- *Unequal Access to Foundational Technology*
- *Amplification of Inequalities*

Concentration of Power & Monopolies of Corporations

Privacy Violation with Intrusive Digital Surveillance:

- *Manipulative Advertising*
- *Behavior Modification*
- *Bystander Privacy*

Alignment with Human Values and Meaning-Making

Unintended Harmful Consequences of AR/VR

- *Filter Bubbles and Echo Chambers*
- *Misinformation*
- *Anti-Social Behavior due to VR Disinhibition*
- *Empathy Used Antisocially*

"Too True to Be Good": Disenchantment with Reality

Neglect of physical self and environment

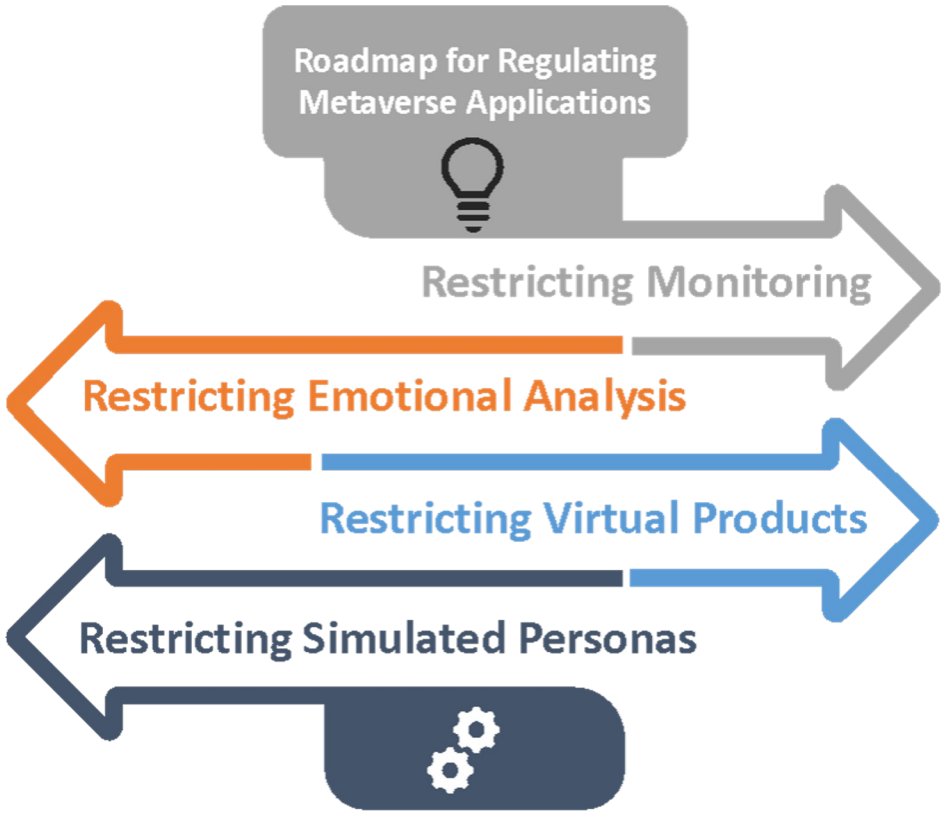
Safety and Security

- *Children Safety*
- *Hacking and Misuse of Identity*
- *Virtual Violence*

Regulating AI-XR Educational Metaverses

Restricting the use of emotional reactions and physiological data for fine-tunning marketing to provide emotion-responsive advertising.

Meatverse patform providers should inform their users about the existance of simulated personas having promotional agenda.



Restricting metaverse service providers to not monitor their users and to not store low-level user data for longer time.

Restricting realistic looking virtual product placement in metaverse for advertisement to avoid targeted-manipulation.

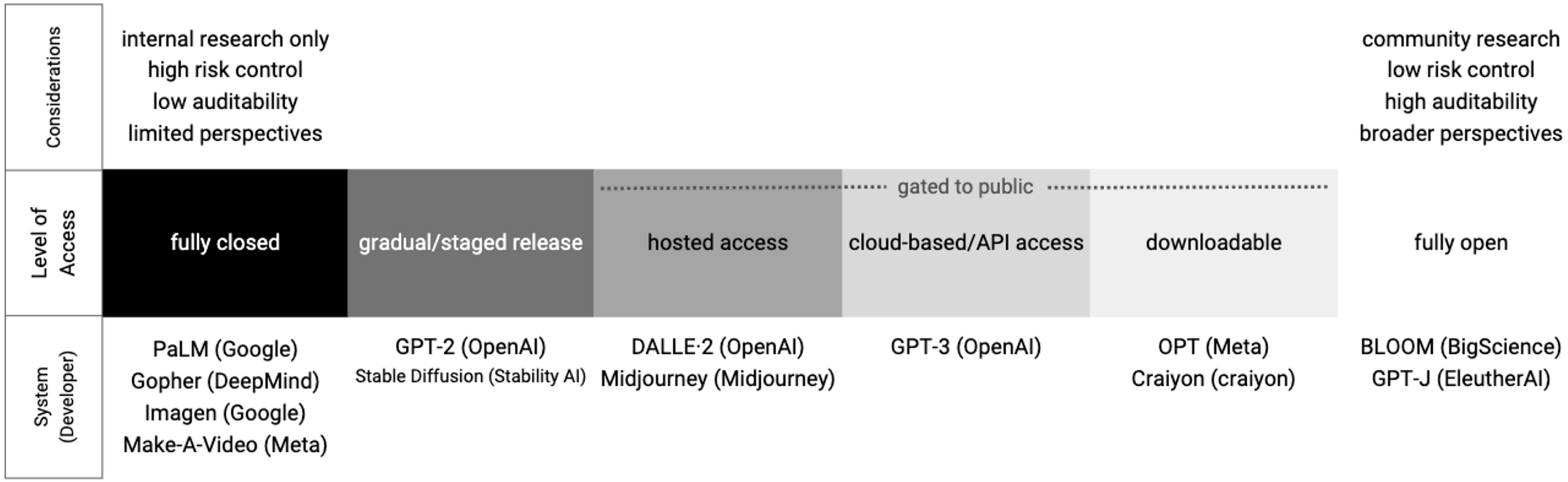
An Introduction to Generative AI (GenAI) and Mixed Reality (MR) and open-source GenAI and MR

GenAI Definition

Generative Artificial Intelligence (GenAI) refers to AI models that are capable of generating new content such as text, images, videos, and music. In recent times, these AI foundation models have been trained on vast amounts of data using self-supervised techniques, training them for multiple downstream tasks, where they generate content often in response to specific prompts.

Types of GenAI models

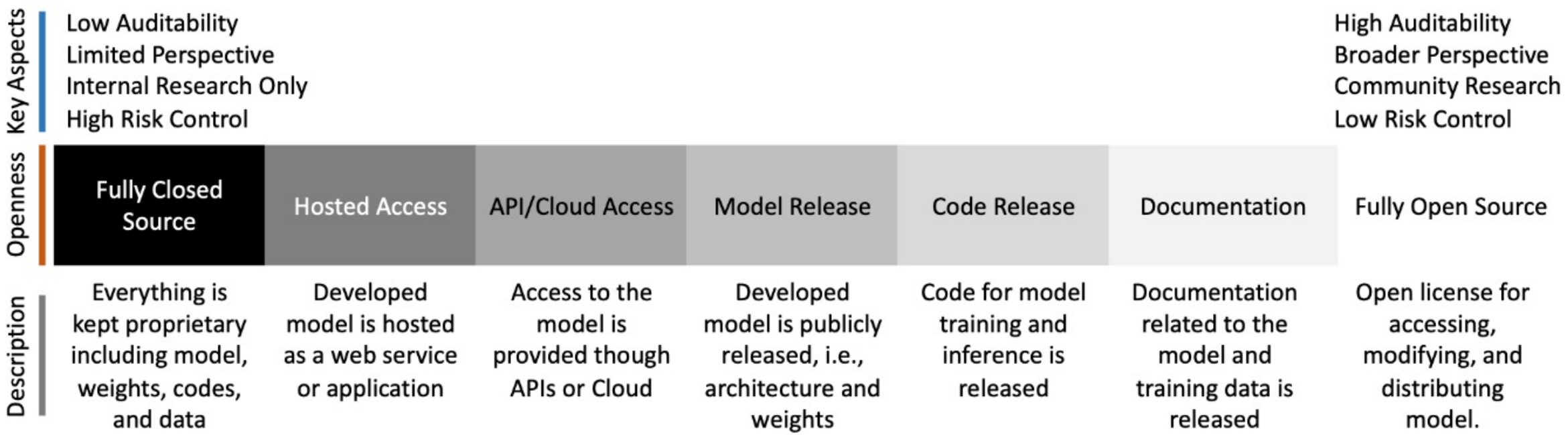
(spectrum of openness in open GenAI models)



Solaiman, I. (2023, June). The gradient of generative AI release: Methods and considerations. In *Proceedings of the 2023 ACM conference on fairness, accountability, and transparency* (pp. 111-122).

Types of GenAI models

(spectrum of openness in open GenAI models)



Types of GenAI models

(spectrum of openness in open GenAI models)

Model Name	Developer	Level of Access	Key Features	Release Strategy
LLaMA 2	Meta	Open-source	Advanced language understanding and generation	Fully open to the public
LLaMA 3	Meta	Open-source	Improved performance and scalability	Fully open to the public
Mistral	Mistral AI Labs	Open-source	Designed for flexibility in various tasks	Fully open to the public
Claude	Anthropic	Hosted access	Safety and alignment-focused	Gated to public, API access
ChatGPT	OpenAI	Hosted access	Conversational AI, context awareness	Gated to public, cloud-based/API

Open-Source MR Platforms

Holokit is a MR platform designed to provide developers with tools for creating immersive experiences by combining AR and VR technologies.

Mozilla Hubs is an open-source platform for creating and sharing virtual reality spaces, allowing users to meet in customizable virtual environments using web browsers or VR headsets.

JanusVR is an open-source immersive internet browser that enables users to browse the web in virtual reality. It allows for the creation and exploration of interconnected virtual spaces.

OpenXR is a cross-platform standard for VR and AR app development. It offers an open-source API, enabling developers to create MR applications compatible with different hardware devices.

A-Frame is an open-source web framework for building VR experiences on the web. It uses HTML-like syntax and works with popular VR platforms like Oculus Rift, HTC Vive, and Windows Mixed Reality.

OSVR is an open-source software platform for virtual and augmented reality. It allows discovery, configuration and operation of several VR/AR devices and peripherals.

Open-Source MR Platforms

Feature/Platform	HoloKit	Mozilla Hubs	JanusVR	OpenXR	A-Frame	OSVR
Type	Mixed Reality	Virtual Reality	Web Browser (VR)	API Standard	Web Framework	Software Platform
Main Use	Immersive apps	Virtual spaces	Immersive browsing	Cross-platform dev.	VR on the web	VR/AR device support
Technology	AR and VR	WebVR	WebVR, HTML5	VR and AR	HTML-like syntax	VR and AR
Device Compatibility	Smartphones	Multiple VR headsets	VR headsets	Various VR/AR	Popular VR headsets	Various VR/AR
Key Features	Affordable, easy	Customizable, social	Portal-like links	Universal API	Easy, accessible	Hardware integration
Accessibility	High	High	Medium	High	High	High
Target Audience	Developers	Users, Developers	Users, Developers	Developers	Developers	Developers

AR Application Development APIs

AR Development Kit	Ownership	Open Source	Platform	Main Features
ARCore	Google	No	Android, iOS	APIs for environmental understanding, motion tracking, and light estimation
ARKit	Apple	No	iOS	Advanced face tracking, realistic rendering, and environmental understanding
ARToolKit+	ARToolworks	Yes	Cross-platform	Marker and markerless tracking, simple to set up for basic AR experiences
Mixare	Independent	Yes	Android	Open-source AR browser and SDK, combines data from multiple sources
Vuforia	PTC Inc.	No	Android, iOS, UWP	Wide range of features including image recognition, 3D object detection, and VuMarks
Wikitude	Wikitude GmbH	No	Android, iOS, Windows	Image recognition, 3D tracking, cloud recognition, and location-based services

GenAI/ Metaverse Case Study:
Interview Training and Education Module (ITEM)

AI-Enhanced Interview Simulation in the Metaverse: Transforming Professional Skills Training through VR and Generative Conversational AI

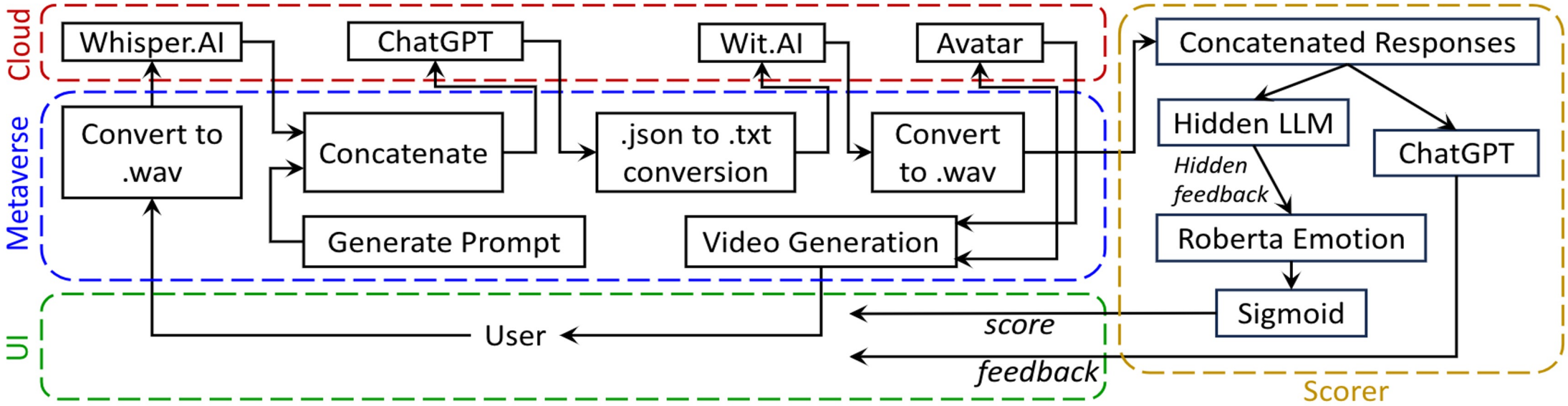
Abdullah Bin Nofal[1], Hassan Ali[1,2], Muhammad Hadi[1], Aizaz Ahmad[1], Adnan Qayyum[1], Aditya Johri[3], Ala Al-Fuqaha[4], and Junaid Qadir[5]

(a) Stage 1: *ITEM* asking first question after initialization.

(b) Stage 2: Student's response to a question is displayed for confirmation.

(c) Stage 3: *ITEM* provides feedback to the student.

GenAI/ Metaverse Case Study:
Interview Training and Education Module (ITEM)



- Utilized C# programming language within the Unity 3D engine.
- Used OpenAI Unity Wrapper for GPT API integration.
- Used Whisper API and Wit.ai for speech-to-text functionality and text-to-speech capabilities.
- To ensure compatibility/performance across VR platforms, we used *OpenXR framework* & *Oculus Integration SDK*.
- We evaluated ITEM using Oculus Quest 2 and Oculus Quest Pro devices.

Some Questions to Consider In Educational Metaverses

- Does the AR/VR project *enhance human engagement* with the physical world and real-world issues?
- *Will the AR/VR amplify existing real-world inequality?*
- Have the *system-level repercussions of AR/VR technology* usage been considered?
- Which policy and regulatory frameworks have been considered to ensure there is not violation of human rights?
- Have the necessary steps been taken to accommodate and promote *diversity and inclusion* (D&I) so that the developed technology is appropriate for the needs and sensitivities of different groups?
- Is there transparency in the system regarding what data is collected, who it is shared with and how it is used?
- What steps and frameworks have been followed to ensure that *any behavior modification or nudging performed by the technology is guided by ethics and law and is culturally sensitive and pro-social*?
- Has appropriate consideration been given to *safeguarding children's rights* if the application is intended for use by children?
- Is the technology *culturally appropriate*?

Conclusions

1. ***Leveraging GenAI and MR for Personalized Learning:***
GenAI and MR technologies enable educational metaverses that can provide dynamic, personalized learning environments that adapt to individual needs and enhance engagement.

2. ***Expanding Educational Accessibility and Inclusivity:***
GenAI and MR can extend access to education through simulations and virtual labs accessible globally, promoting inclusivity for students in remote or underserved areas.

3. ***Promoting Transparent Innovation with Open-Source Technologies:***
Open-source GenAI/MR can foster collaborative development, enabling rapid innovation, and improved security through community involvement.

Thank you!

jqadir@qu.edu.qa