Citiverse Initiative

Global Initiative on Virtual Worlds and AI – Discovering the Citiverse

No One Left Behind in the Citiverse: A Blueprint for Accessible AI-Powered Virtual Worlds

Foreword

This publication was developed within the framework of the Global Initiative on Virtual Worlds and AI – *Discovering the Citiverse*, which is a global multistakeholder platform launched by the International Telecommunication Union (ITU), the United Nations International Computing Centre (UNICC), and Digital Dubai, and supported by more than 70 international partners.

The Initiative aims to shape a future where AI-powered virtual worlds are inclusive, trusted, and interoperable. By connecting people, cities, and technologies, it empowers meaningful progress through AI-powered virtual worlds.

Acknowledgements

The development of this deliverable was led and coordinated by Pilar Orero (Universitat Autònoma de Barcelona (UAB), Spain).

This publication was authored by Pilar Orero (UAB, Spain), Paola Cecchi Dimeglio (Harvard University), Puneet Jain (Concordia University, Montreal, Canada Zurich University of the Arts, Switzerland), Krzysztof Krejtz (SWPS University, Poland), Yong Jick Lee (Center for Accessible ICT, Korea (Rep. of), Sarah Anne McDonagh (UAB, Spain), Alessandro Pollini (Università Telematica Internazionale Uninettuno, Italy), Anna Matamala (UAB, Spain), Estella Oncins (UAB, Spain), Chris J Hughes (Salford University, United Kingdom), Carlos Duarte (LASIGE, Faculdade de Ciências da Universidade de Lisboa, Portugal).

The authors extend their sincere thanks to the Executive Committee of the Global Initiative on Virtual Worlds and AI: H.E. Mr Hamad Al Mansoori (Director General, Digital Dubai), H.E. Mr Jerry William Silaa (Ministry of Information, Communication and Information Technology, Tanzania), H.E. Mr William Kabogo Gitau (Ministry of Information, Communications and the Digital Economy, Kenya), Felipe Fernando Macías Olvera (Municipality of Queretaro, Mexico), Manuel Barreiro (Aston Group), Karl-Filip Coenegrachts (Open & Agile Smart Cities (OASC)), Hyoung Jun Kim (ITU-T Study Group 20 "Internet of Things, digital twins and smart sustainable cities and communities"), Jaakko Mustakallio (City of Tampere, Finland), Paula Llobet Vilarrasa (City of Valencia, Spain), Sameer Chauhan (United Nations International Computing Centre (UNICC)), and Jeong Kee Kim (World Smart Sustainable Cities Organization (WeGO)).

The authors also thank the Steering Committee of the Global Initiative on Virtual Worlds and AI for their continued support: Okan Geray (Dubai Digital Authority), Bertrand Levy (The Sandbox), Teppo Rantanen (City of Tampere, Finland), Paola Cecchi Dimeglio (Harvard University), Ernesto Faubel (European Digital Infrastructure Consortium (EDIC) on Local Digital Twins), Martin Brynskov (OASC), Anish Sethi (UNICC), AnaMaria Meshkurti (AMVS Capital) and Roland van der Heijden (City of Rotterdam, The Netherlands).

The authors also extend their gratitude to the contributing organizations along with their representatives: Cristina Bueti, Yining Zhao, Chiara Co (ITU) and Franca Vinci (UNICC).

Disclaimers

The opinions expressed in this publication are those of the authors and do not necessarily represent the views of their respective organizations, Executive Committee members or Steering Committee members of the Initiative. The findings presented in this report are based on a comprehensive review of existing literature and voluntary written contributions submitted by a diverse range of stakeholders.

ISBN

978-92-61-41471-9 (Electronic version)
978-92-61-41951-6 (EPUB version)

No One Left Behind in the Citiverse: A Blueprint for Accessible AI-Powered Virtual Worlds

Table of contents

List of tables

Tables

Abbreviations and acronyms

AI	Artificial Intelligence
AR	Augmented Reality
AMT	Advanced Manufacturing Technologies
BCI	Brain-Computer Interface
CNC	Computerized Numerical Control
ECG	Electrocardiography
EEG	Electroencephalography
EMG	Electromyography
EPUB	Electronic Publication
GDPR	General Data Protection Regulation
HCI	Human-Computer Interaction
ICT	Information and Communication Technology
IM	Industrial Metaverse
IMU	Inertial Measurement Unit
ISO	International Organization for Standardization
ITU	International Telecommunication Union
LMS	Learning Management System
LLM	Large Language Model
MR	Mixed Reality
NLP	Natural Language Processing
OASC	Open & Agile Smart Cities
PPG	Photoplethysmography
RBAC	Role-Based Access Control
RGB	Red, Green, Blue
RTP	Real-time Transport Protocol
SMA	Spinal Muscular Atrophy
SDH	Subtitles for the Deaf and Hard-of-Hearing
SME	Small and Medium-Sized Enterprises

(continued)

UNICC	United Nations International Computing Centre
VR	Virtual Reality
VRML	Virtual Reality Modeling Language
WAI	Web Accessibility Initiative
WCAG	Web Content Accessibility Guidelines
W3C	World Wide Web Consortium
WebAuthn	Web Authentication
WebXR	Web-based Extended Reality
XR	Extended Reality
XAUR	XR Accessibility User Requirements
X3D	Extensible 3D

Executive summary

This report underscores the pressing need for digital inclusion and accessibility in our rapidly evolving digital landscape, particularly in a cityscape where a pervasive digital layer intersects with physical urban environments. As cities globally integrate smart infrastructure with immersive technologies, the fundamental aspects of people's lives, including commuting, work and access to public services, are undergoing a profound transformation. At this pivot point of technological and societal change, we strive to make every effort to ensure the accessibility by design of this emerging world, thereby ensuring that no one is left behind, irrespective of their age, capabilities, social status, digital literacy, or disabilities. The equitable access in virtual worlds is not just a technical challenge, but a fundamental social imperative. We need to ensure that accessibility also involves addressing critical ethical considerations and avoiding bias. This might be achieved by fostering a genuine sense of presence and co-presence for diverse potential users.

As immersive and AI-powered technologies reshape the fabric of urban life, the need for inclusive and accessible digital environments has become more urgent than ever. Virtual worlds, when thoughtfully designed, have the potential to enhance education, public services, cultural participation, and economic opportunities. Without proactive measures, these same technologies risk reinforcing existing barriers.

This report titled 'No One Left Behind in the Citiverse: A Blueprint for Accessible AI-Powered Virtual Worlds' developed under the Global Initiative on Virtual Worlds and AI – *Discovering the Citiverse*, highlights how digital transformation can, and should, be made inclusive from the outset. The publication brings together insights from across sectors and disciplines to present a forward-looking framework for ensuring that everyone, regardless of age, capability, or digital literacy, can participate fully in emerging digital ecosystems.

The report introduces actionable models, practical use cases, and collaborative approaches to integrate accessibility across the lifecycle of virtual world development. It also underscores the value of common standards to promote seamless, user-friendly experiences across platforms and jurisdictions.

Overview of the Chapters

Chapter 1: Introduction Frames the challenge and opportunity of digital inclusion in the context of smart cities and virtual environments, emphasizing the societal and technological shifts underway.

Chapter 2: The Four Pillars of Stakeholder Engagement Outlines the essential roles of the public sector, private sector, civil society, and academic and standards bodies in driving accessible innovation and governance.

Chapter 3: Framework for Virtual Worlds Interaction Proposes a foundational model of modular "atomic interactions" (e.g., navigation, communication, authentication), enabling scalable, interoperable, and inclusive user experiences.

Chapter 4: Tools, Terminologies and Technologies in XR Maps out key technologies, from hand tracking and haptics to brain-computer interfaces, and their relevance for designing user-centric, accessible environments.

Chapter 5: Accessibility Services Details a wide range of services, including subtitles, audio description, sign language, haptic cues, and simplified language, essential to making immersive environments usable by diverse populations.

Chapter 6: Use Cases Showcases global examples where accessibility has been successfully implemented in virtual environments, such as immersive corporate training, education, accessible tourism, industrial skills development, and inclusive digital art.

Chapter 7: Standardization Reviews the current landscape of international standardization efforts and highlights the importance of coordination to avoid fragmentation and ensure global interoperability.

Chapter 8: Conclusions Presents practical guidance for policymakers, developers, and standards bodies to scale accessibility in virtual worlds, through shared frameworks, inclusive design principles, and cross-sector collaboration.

This publication demonstrates that inclusion is not just a technical add-on, it is a strategic enabler of innovation, resilience, and global reach. As governments and organizations explore the next frontier of digital transformation, ensuring accessibility by design will be critical to building sustainable, people-centred virtual futures.

1 Introduction

This report underscores the pressing need for digital inclusion and accessibility in our rapidly evolving digital landscape, particularly in a cityscape where a pervasive digital layer intersects with physical urban environments. As cities globally integrate smart infrastructure with immersive technologies, the fundamental aspects of people's lives, including commuting, work and access to public services, are undergoing a profound transformation. At this pivotal moment of technological and societal transformation, it is essential to embed accessibility by design into the foundations of emerging virtual worlds – ensuring that no one is excluded, regardless of age, ability, social status, or digital literacy. The equitable access in virtual worlds is not just a technical challenge, but a fundamental social imperative. We need to ensure that accessibility also involves addressing critical ethical considerations and avoiding bias. This might be achieved by fostering a genuine sense of presence and co-presence for diverse potential users.

To address these challenges, the report emphasizes the importance of robust accessibility standards that provide a structured foundation for designing interactions within virtual worlds, promoting modularity, inherent accessibility by design, and seamless interoperability. The report is a result of a collective effort aiming to ensure that emerging technologies, particularly those integral to smart cities and artificial intelligence, are designed to prevent exclusion, promoting better mobility and equitable access to information for all. The key contributions towards this goal are as follows.

The virtual worlds, as presented in this document, are envisioned as a transformative digital layer that integrates with physical urban environments, creating a "citiverse". It is a space in which artificial intelligence (AI) becomes human-centred, converting data into dynamic and immersive experiences. The document emphasizes that accessibility is fundamental to ensuring universal participation in this evolving era of digital transformation, highlighting its role in enabling equitable access and avoiding bias within virtual worlds.

Firstly, the report defines *Four Pillars of Stakeholder Engagement*, identifying the public sector, private sector, communities, academic and standardization bodies as essential collaborators for building inclusive virtual spaces. Secondly, it introduces a comprehensive *Framework for Virtual Worlds Interaction*, which standardizes core (atomic) interactions (digital actions) like authentication, communication and navigation to ensure built-in accessibility. Thirdly, the report provides an overview of *Input/Output in Extended Reality (XR) – Tools and Technologies*, covering diverse interaction methods from hand and eye tracking to bio-signal based interfaces and the integration of existing assistive technologies like white canes and wheelchairs into XR experiences. Fourthly, it categorizes and explains various *Accessibility Services* for virtual and extended reality such as textual (subtitles, transcripts), audio (audio description, audio introductions, audio subtitles), visual-gestural (sign language interpreting) and haptic alternatives, along with the crucial role of easy-to-understand languages (Easy Language, Plain Language) for accommodating diverse users.

Furthermore, the Report presents compelling *Use Cases* demonstrating all the key elements of the human-centric design and the practical benefits of digital inclusion in virtual settings. These encompass various aspects of human activity, including work, leisure, sport and art. The use cases showcase accessible learning platforms, immersive corporate training, accessible cultural heritage, an accessible virtual reality (VR) boxing platform tailored for blind users and innovative mouth-based interaction for artists with motor disabilities. Lastly, the report provides an essential overview of *Standardization* efforts by key international bodies, highlighting their contributions and the growing convergence towards comprehensive digital inclusion through a unified approach.

2 The four pillars of stakeholder engagement

As cities integrate smart infrastructure with immersive technologies to create a "citiverse" – a digital layer over physical urban life – the need for inclusion and accessibility becomes urgent. This transformation affects how people live, move, work and access public services. To ensure no one is left behind, four key categories of stakeholders must work together: the public sector, private sector, communities and a vital fourth group – academic and standardization stakeholders.

2.1 Public sector and governance stakeholders

City governments, urban planners, public agencies and national regulators set the vision, policies and frameworks that guide citiverse development. Their role is to ensure that public investments and digital infrastructure are designed with accessibility in mind from the outset. From mandating that digital city services meet accessibility standards to investing in broadband equity and public digital literacy programmes, the public sector acts as regulator and enabler. It also plays a coordinating role, fostering collaboration among stakeholders and ensuring that inclusion and accessibility are central to smart city governance.

2.2 Private Sector and Technology Providers

Technology companies, telecom providers, companies of different sizes and maturity, start-ups and infrastructure firms drive innovation in the citiverse. They build the platforms, software and connectivity that underpin virtual city experiences. But with this power comes opportunities and responsibility: the private sector must design products and services with diverse users in mind – across ability, gender, race, ethnicity, age, language, income, and more. From incorporating accessibility features into virtual reality platforms to ensuring affordable access to communication technologies, tech developers must embed universal design principles at every stage. They should also take the opportunity to invest in developing accessibility since they will capitalize on innovation. When aligned with public goals and user needs, the private sector can become a powerful engine for inclusive and accessible digital transformation.

2.3 Community and civil society stakeholders

Local residents, grassroots organizations, advocacy groups and non-profits ensure that the citiverse remains people-centred. These stakeholders offer critical insights into the lived experiences of underserved or marginalized populations. Their voices help shape services that reflect users' needs – whether by advocating for multilingual interfaces, co-designing disability-friendly features, or highlighting gaps in digital access. Inclusion efforts fail when they are imposed top-down; communities must be active partners in citiverse design and governance. Civil society also plays a watchdog role, holding governments and companies accountable for their accessibility and equity commitments.

2.4 Academic and standardization stakeholders

The fourth pillar comprises researchers, universities, think tanks and standards developing organizations (SDOs). These knowledge partners bring evidence-based insights, helping cities and developers understand how different users engage with emerging technologies. They develop guidelines, accessibility benchmarks and technical standards to ensure systems are inclusive and interoperable. Institutions like International Telecommunication Union (ITU) and World Wide Web Consortium (W3C), for example, set global standards that make digital services usable across devices and abilities. Academic partners also train future leaders in inclusive design and advice on ethical, data-driven policymaking. Their independent perspective and long-term focus help embed inclusion and accessibility as foundational values, not as afterthoughts.

Together, these four stakeholder categories form the backbone of a robust and inclusive governance ecosystem for the citiverse. Each plays a distinct and indispensable role, and lasting success will depend on their sustained collaboration, real-time feedback loops, and an unwavering collective commitment to digital equity and universal access.

3 Framework for virtual worlds interaction

3.1 Introduction

Accessibility in virtual worlds extends beyond content presentation to include navigational structures, interaction mechanisms, and the integration of assistive technologies. Addressing these challenges requires a systematic and standardized approach to interaction design that accommodates diverse user needs while remaining scalable across platforms and use cases.

The framework for Virtual Worlds Interaction plays a critical role in promoting consistency, re-usability, and inclusivity. Positioned between low-level services and high-level use cases, the framework defines modular atomic interactions such as authentication, communication and object manipulation that can be combined flexibly to support a wide range of user goals.

By abstracting common interaction patterns and integrating accessibility and interoperability principles at their core, the framework enables developers to build inclusive applications without reinventing foundational elements. It also supports standardization efforts by providing a common vocabulary and structure for designing and evaluating virtual experiences. As such, the framework for Virtual Worlds Interaction is an essential building block in the effort to create equitable, sustainable and user-centred virtual environments.

3.2 Objective of the framework

The primary objective of this framework is to provide a structured foundation for standardizing interactions within virtual worlds, emphasizing modularity, accessibility and interoperability. Modularity ensures that interactions are re-usable and adaptable across various contexts, simplifying implementation and maintenance. Accessibility aims to guarantee equitable use of virtual environments by users with diverse abilities, addressing barriers from initial design stages. Interoperability facilitates seamless integration of different technological components and services, allowing frameworks to function consistently across multiple virtual platforms and scenarios.

3.3 Components of the framework

To achieve these objectives, the framework comprises three core components: atomic interactions, support services and user-centric design.

3.3.1 Atomic interactions

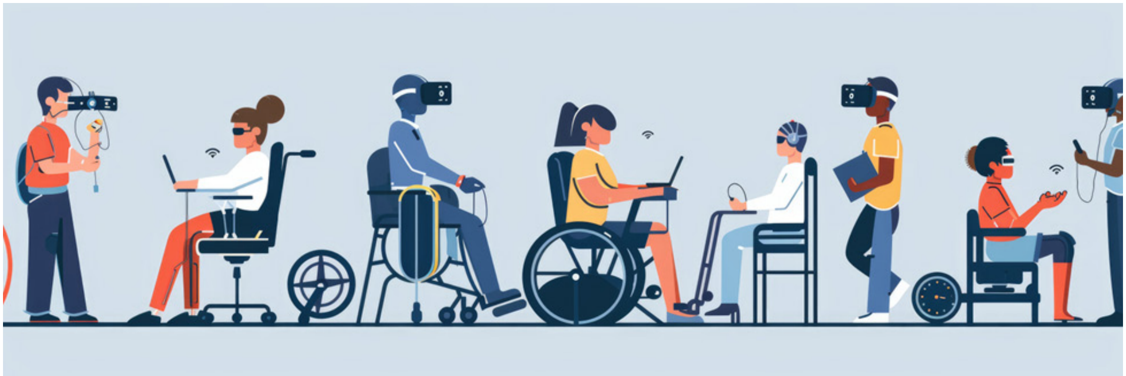
Atomic interactions are fundamental interaction patterns, serving as re-usable building blocks essential for constructing complex user experiences in virtual worlds. By abstracting common tasks such as authentication, navigation and communication, atomic interactions provide standardized protocols that can be consistently applied across diverse virtual environments, simplifying development and promoting consistency and usability.

3.3.2 Support services

Support services encompass the technological resources and capabilities that facilitate atomic interactions, enhancing their functionality and accessibility. Examples include subtitling services, real-time language translation, assistive navigation aids and text-to-speech systems. By providing these enabling technologies, support services extend the functionality of atomic interactions, ensuring they meet diverse user needs and preferences.

3.3.3 User-centric design

User-centric design underpins the entire framework by emphasizing the principles of accessibility, usability and inclusivity throughout the interaction design process. This involves proactively considering diverse user requirements such as sensory, cognitive and motor abilities, from the earliest design stages. Adhering to established accessibility standards such as the Web Content Accessibility Guidelines (WCAG)[1] ensures that the framework not only supports universal usability but also aligns with legal and ethical obligations toward digital inclusion.



3.4 Atomic interactions

3.4.1 Authentication

Definition and role

Authentication is the process of verifying the identity of users entering virtual environments. Its primary role is to establish trust, secure user data and ensure that only authorized individuals gain access to virtual spaces and services. Reliable authentication safeguards user privacy and provides a foundation for secure interactions within the virtual environment. For example, authentication allows users to securely log into a virtual workspace or perform sensitive actions such as accessing confidential documents.

Design and implementation

Effective authentication mechanisms in virtual worlds should emphasize usability, security and accessibility. User-centric designs should accommodate diverse user needs, including cognitive, sensory and motor abilities. Implementations can range from traditional methods (username/password), biometrics (e.g., facial recognition, fingerprint scans), to multifactor authentication, balancing convenience and security. Support services such as screen readers, voice authentication and haptic feedback can enhance accessibility, so ensuring that authentication processes are inclusive and usable for all users.

Standards and guidelines

Authentication in virtual environments should comply with recognized industry standards and accessibility guidelines. Relevant standards include the ISO/IEC 27001[2] for information security management, OAuth 2.0[3] for authorization protocols and Web Authentication (WebAuthn)[4] specifications from the W3C. Additionally, design and implementation must adhere to the WCAG to support universal access, particularly guidelines addressing keyboard accessibility, error prevention and understandable authentication procedures.

3.4.2 Role and permission management

Definition and role

Role and permission management involves assigning, defining and enforcing user roles and their associated permissions within virtual environments. It is essential for maintaining security, ensuring appropriate access to resources and facilitating effective collaboration. Proper management prevents unauthorized access and protects sensitive information, aligning virtual interactions with organizational policies and regulatory requirements. For example, role and permission management ensures that only authorized personnel can access secure virtual rooms such as confidential meeting spaces or restricted data repositories.

Design and implementation

Implementing role and permission management requires clear definitions of user roles and associated permissions tailored to user responsibilities and tasks. Designs should emphasize intuitive and accessible user interfaces for managing permissions, ensuring usability for administrators and users with varying abilities. Support services may include simplified administrative dashboards, auditory confirmations and accessible notifications that communicate permission status clearly. User-centric design principles ensure that role assignments and permission requests are transparent, understandable and accessible for all users, accommodating diverse cognitive and sensory abilities.

Standards and guidelines

Role and permission management systems should adhere to established standards such as the Role-Based Access Control (RBAC) model outlined in NIST SP 800-162.[5] RBAC provides a structured approach for defining roles, permissions and user assignments, enhancing security and manageability. Additional relevant standards include ISO/IEC 27001:2022[6] for information security management and ISO/IEC 29146:2024,[7] which provides comprehensive guidelines for managing access permissions securely and consistently. Adherence to the Web Content Accessibility Guidelines ensures that interfaces for role and permission management remain accessible and usable for all users, addressing requirements related to operability, comprehensibility and assistive technology compatibility.

3.4.3 Navigation

Definition and role

Navigation involves enabling users to orient themselves and move efficiently within or between virtual spaces. Its role is critical in providing spatial awareness, supporting users in locating points of interest and facilitating smooth transitions across virtual environments. For example, navigation allows users to find a specific meeting room within a virtual office or explore various stores within a virtual shopping mall.

Design and implementation

Effective navigation systems must be intuitive, accessible and responsive to diverse user abilities. Design strategies include clear visual cues, audio prompts and tactile feedback to support orientation and movement. Implementation may involve interactive maps, wayfinding aids, teleportation mechanisms or continuous locomotion. Support services such as descriptive audio guides, text-to-speech systems and personalized assistance ensure that navigation is accessible and inclusive, so accommodating users with visual, auditory or motor impairments.

Standards and guidelines

Navigation in virtual environments should adhere to established ergonomic and usability standards. Relevant standards include ISO/IEC 18040:2019,[8] defining reference models and interactions for live actors and entities in mixed and augmented reality; ISO/IEC 18039,[9] providing a comprehensive reference model for mixed and augmented reality systems; and ISO 9241-394:2020,[10] covering ergonomic principles for virtual reality navigation and interaction. Additionally, IEEE Std 3079-2020[11] provides usability guidelines for head-mounted display systems and W3C Immersive Web[12] (WebXR), provides standards for navigation and interaction in web-based VR and AR environments. Navigation systems should also comply with the WCAG to ensure universal accessibility and usability.

3.4.4 Proprioception

Definition and role

Proprioception in virtual environments refers to a user's ability to perceive the position, movement and orientation of their body or avatar within the virtual space. It is essential for achieving embodied presence, improving spatial awareness and enabling natural interaction with virtual objects and environments. Accurate proprioceptive feedback allows users to navigate, interact and respond within the environment with increased confidence and reduced cognitive load. For example, proprioception enables a user to reach out and grasp a virtual object with realistic alignment between their physical and virtual hand positions.

Design and implementation

Designing for proprioception involves aligning virtual feedback with users' physical movements to reinforce embodiment and reduce disorientation. This includes accurate avatar representation, synchronized motion tracking, and appropriate visual, auditory and haptic feedback. Implementation may involve the use of motion capture systems, inertial measurement units (IMUs), depth sensors, or wearable haptic devices. Support services include avatar customization, real-time body tracking middleware and calibration tools to adjust for individual body dimensions or mobility constraints. User-centric design must accommodate a range of physical abilities and comfort levels, offering options to reduce motion sickness, support seated or limited-mobility use and allow users to adjust the pace and fidelity of proprioceptive feedback.

Standards and guidelines

Proprioception in virtual environments should follow established ergonomic and interaction design standards. ISO 9241-394:2020[13] provides principles for immersive virtual environments, including body tracking and feedback alignment. IEEE Std 3079-2020[14] offers guidelines for the usability of head-mounted displays, covering proprioceptive aspects of head and hand movement. Additionally, ISO/IEC 23090-2:2023[15] (Omnidirectional media format (OMAF)) supports omnidirectional media experiences, which are relevant for maintaining spatial orientation. Conformance with accessibility guidelines such as WCAG and ISO/IEC 24751[16] (individualized adaptability), ensures that proprioceptive feedback mechanisms are usable and customizable for users with diverse physical and perceptual needs.

3.4.5 Object manipulation

Definition and role

Object manipulation refers to the interaction techniques that allow users to select, move, transform and otherwise interact with virtual objects within a virtual environment. It plays a central role in enabling direct engagement with the environment, fostering a sense of agency and control. This interaction is foundational for tasks such as assembling virtual prototypes, organizing virtual workspaces or simulating real-world tasks. For example, a user might rearrange digital documents on a virtual desk or construct a model using 3D blocks in a collaborative design scenario.

Design and implementation

Designing for object manipulation requires a balance between precision, intuitiveness and accessibility. Interaction models may include direct manipulation (e.g., grabbing with motion-tracked hands), indirect control (e.g., pointer-based systems), or gesture recognition. Implementation involves the integration of hand tracking, collision detection and physics simulation to provide realistic and responsive object behaviours. Support services such as haptic feedback, audio cues and visual highlights enhance the user's ability to perceive object state and affordances. User-centric design must ensure that object manipulation is usable for individuals with limited dexterity or alternative input needs, providing features like configurable control sensitivity, snap-to-grid alignment and voice-activated manipulation options.

Standards and guidelines

Object manipulation in virtual environments should align with standards that address input ergonomics and interactive system behaviour. ISO 9241-940:2017[17] offers guidance on tactile and haptic interactions, including feedback mechanisms relevant for manipulation tasks. ISO 9241-411:2012[18] provides requirements for input devices and interaction techniques, ensuring usability and ergonomic soundness. ISO/IEC 14772[19] (Virtual Reality Modeling Language (VRML)) and ISO/IEC 19775[20] (X3D) specify foundational technologies for describing interactive 3D objects. Adherence to the WCAG and ISO/IEC 24751[21] promotes accessibility through adaptable interfaces, so ensuring object manipulation is inclusive across varying user abilities and preferences.

3.4.6 Communication

Definition and role

Communication in virtual environments refers to the exchange of information between users, avatars or system agents through various modalities such as voice, text, gesture and visual signals. It is essential for collaboration, social interaction and task coordination, supporting synchronous and asynchronous exchanges. Effective communication enables users to interact naturally and meaningfully, so fostering immersion and trust. For example, users in a virtual meeting room may use spatial voice chat and hand gestures to discuss project updates and share virtual documents in real time.

Design and implementation

Communication systems in virtual worlds must support multimodal interaction while maintaining clarity, reliability and inclusivity. Design should accommodate different communication preferences and abilities, offering configurable channels such as text-to-speech, speech-to-text, sign language avatars and real-time language translation. Implementation includes networking protocols for real-time media exchange, synchronization of voice and gestures and spatial audio rendering. Support services may include subtitling engines, emotion recognition and chat moderation tools to enhance communication quality and safety. User-centric design must prioritize ease of access, customizability and minimal cognitive load, enabling users with sensory, cognitive, or linguistic challenges to participate fully in conversations.

Standards and guidelines

Communication systems in virtual environments should align with interoperability and accessibility standards. W3C WebRTC[22] provides the foundation for real-time audio and video communication over the web. IETF RFC 3550[23] defines the Real-time Transport Protocol (RTP), essential for low-latency media streaming. ISO/IEC 20006[24] addresses interoperability of language resources, supporting multilingual communication. For accessibility, WCAG 2.2 and ISO/IEC 24751[25] provide principles for accessible communication interfaces, including captioning, alternative input and assistive presentation formats. Conformance with these standards ensures communication is effective, secure and inclusive across diverse user populations.

3.4.7 Consent management

Definition and role

Consent management in virtual environments refers to the processes through which users grant, deny or revoke permission for specific actions, data processing or interactions. It is a critical mechanism for upholding user autonomy, privacy and trust in immersive systems. This interaction ensures that users maintain control over personal data, participation in shared spaces and exposure to sensitive content or interactions. For example, a user may be prompted to give consent before their biometric data are collected for avatar personalization or before entering a virtual area where recordings are taking place.

Design and implementation

Effective consent management must prioritize clarity, transparency and reversibility. Interfaces should explain clearly what is being requested, why, how long consent is valid and how users can change their decisions. Design should employ plain language, consistent visual indicators and accessible formats to ensure all users understand their choices. Implementation may include consent dashboards, contextual prompts and logging mechanisms for auditing and compliance. Support services such as screen readers, voice-based confirmations and translation tools ensure users of all abilities and languages can navigate consent interactions. User-centric design requires minimizing cognitive load and offering granular controls over specific permissions, ensuring users retain meaningful agency.

Standards and guidelines

Consent management should align with data protection and accessibility standards. The EU General Data Protection Regulation (GDPR)[26] and the ISO/IEC 29184:2020[27] standard for online privacy notices and consent provide foundational principles for transparency, informed consent and data handling practices. ISO/IEC 27560:2023[28] further specifies requirements for consent recordkeeping. Accessibility considerations are covered under WCAG 2.2 and ISO/IEC 24751,[29] which support adaptable and comprehensible consent interfaces for users with diverse needs. These standards collectively ensure that consent mechanisms in virtual environments are lawful, ethical and inclusive.

3.4.8 Environmental customization

Definition and role

Environmental customization refers to the user's ability to adjust elements of the virtual environment to match personal preferences, needs or situational requirements. This interaction enhances comfort, usability and accessibility, enabling users to shape their experiences for better cognitive, sensory or emotional alignment. For example, a user with light sensitivity may adjust the brightness or contrast of a virtual workspace, while another might modify ambient sounds or layout configurations for improved concentration and ease of navigation.

Design and implementation

Environmental customization should support a wide range of modifiable parameters such as lighting, colour schemes, audio levels, spatial layout and interaction pace. Design must ensure that customization options are easy to locate, understand and modify. Interfaces should offer pre-set accessibility modes (e.g., high contrast, quiet mode) and allow fine-grained control for advanced users. Implementation includes settings menus, context-aware suggestions and real-time feedback to confirm changes. Support services like profile-based personalization, adaptive systems and assistive configuration tools help users tailor environments dynamically. A user-centric approach emphasizes consistency, reversibility and inclusivity, ensuring users with varied abilities can access and benefit from customization features.

Standards and guidelines

Environmental customization should align with established accessibility and usability standards. WCAG 2.2 provides key guidelines for adaptable content presentation, including requirements for user control over visual and auditory elements. ISO/IEC 24751[30] supports individualized user preferences in digital learning and information environments, directly applicable to virtual customization. ISO 9241-112:2017[31] outlines ergonomic principles for information presentation, ensuring that environment adjustments remain effective and comprehensible. Compliance with these standards guarantees that customization features are functional and inclusive, supporting equitable user experiences in virtual worlds.

3.4.9 Contextual help and guidance

Definition and role

Contextual help and guidance refer to the provision of real-time, situational support to users as they interact with a virtual environment. This interaction helps users understand system functions, resolve uncertainties and perform tasks more efficiently. It plays a key role in onboarding, task execution and recovery from errors. For example, a user navigating a virtual municipal service centre may receive step-by-step prompts guiding them through the process of applying for a digital ID card.

Design and implementation

Effective contextual help must be timely, relevant and unobtrusive. Design should offer multiple modalities – such as visual tooltips, audio narration, or guided tours – adapted to user preferences and abilities. Implementation may include intelligent agents, overlays, help buttons and gesture-triggered assistance that respond to user behaviour and environmental context. Support services like natural language processing (NLP), machine learning-based recommendation systems and multilingual content delivery enable dynamic, personalized help. A user-centric approach ensures that guidance systems are not overwhelming, allow user control over the level of assistance and integrate with assistive technologies to support users with diverse cognitive and sensory needs.

Standards and guidelines

Contextual help and guidance systems should follow usability and accessibility standards to ensure effective and inclusive support. ISO 9241-112:2017[32] provides ergonomic principles for presenting information in interactive systems, including context-aware information delivery. ISO 9241-11:2018[33] outlines usability guidelines and user assistance principles. WCAG 2.2 ensures that help mechanisms are perceivable, operable and understandable across a range of user abilities. Additionally, ISO/IEC 24751[34] supports individualized learner preferences and can guide the design of personalized support systems in virtual environments. These standards ensure that contextual help remains accessible, efficient and respectful of user autonomy.

3.4.10 Feedback and interaction confirmation

Definition and role

Feedback and interaction confirmation refer to the system's responses that inform users about the outcome of their actions within a virtual environment. These responses are essential for maintaining user awareness, confidence and trust in the system, ensuring that interactions are recognized and appropriately processed. For example, when a user submits a virtual form, they may receive an auditory chime, visual confirmation message and haptic pulse confirming the action was successful.

Design and implementation

Effective feedback should be immediate, clear and appropriately matched to the context and user expectations. Design strategies should include multiple feedback modalities – visual (e.g., highlights, animations), auditory (e.g., sounds, spoken cues) and haptic (e.g., vibration, force feedback) – to accommodate user preferences and abilities. Implementation requires consistent response timing, clear status indicators and mechanisms to confirm successful and unsuccessful interactions. Support services such as customizable feedback settings, notification systems and assistive technologies (e.g., screen readers or alternative input devices) enhance accessibility. User-centric design ensures that feedback is non-intrusive, easy to interpret and adjustable for different sensory and cognitive needs.

Standards and guidelines

Feedback and confirmation mechanisms should align with standards for usability, accessibility and user interface design. ISO 9241-210:2019[35] outlines human-centred design principles, including effective feedback provision. ISO 9241-112:2017[36] focuses on information presentation, particularly clarity and consistency of system responses. WCAG 2.2 provides accessibility guidance for perceivable and understandable notifications and status messages, while ISO/IEC 24751[37] supports adaptable feedback presentation tailored to user preferences. These standards ensure feedback mechanisms are inclusive, reliable and supportive of diverse interaction needs.

3.4.11 Session persistence and state management

Definition and role

Session persistence and state management refer to the system's ability to maintain user context, preferences and progress across sessions in a virtual environment. This interaction ensures continuity of experience, enabling users to resume tasks seamlessly, preserve personalized settings and avoid redundant actions. For example, a user working in a virtual design studio can return to their workspace to find tools, objects and layouts exactly as they left them, enhancing productivity and user satisfaction.

Design and implementation

Designing for session persistence requires a robust architecture for saving, retrieving and synchronizing user states across devices and time. This includes tracking environment settings, user interactions, object positions and system preferences. Implementation strategies may involve cloud-based storage, local caching, version control and checkpointing mechanisms. Support services such as user profiles, auto-save features, synchronization APIs and recovery tools enable reliable session management. A user-centric design approach emphasizes transparency (e.g., clear save indicators), user control (e.g., manual save points) and adaptability (e.g., device-agnostic state restoration), ensuring that all users, including those with cognitive or memory impairments, can benefit from consistent and personalized experiences.

Standards and guidelines

Session persistence and state management should follow standards for system interoperability, data integrity and usability. ISO/IEC 23026:2023[38] offers guidelines for website engineering and content management, including session control and personalization principles applicable to virtual environments. ISO/IEC 25010:2023[39] defines quality attributes such as reliability, recoverability and operability essential for stateful systems. WCAG 2.2 and ISO/IEC 24751[40] support accessible mechanisms for saving and resuming interactions, ensuring these features accommodate diverse user needs.

3.5 Framework interoperability and integration

The Framework for Virtual Worlds Interaction functions as an intermediary layer between abstract interaction models and concrete technical implementations. Atomic interactions such as authentication, communication, or object manipulation, depend on a variety of enabling services to operate effectively. These include but are not limited to identity management systems, real-time communication protocols and accessibility tools (e.g., screen readers, speech recognition). By modularizing these services, the framework ensures that interaction logic remains consistent while allowing services to be updated, replaced or extended without disrupting the user experience. This separation of concerns supports scalability, technical flexibility and platform independence.

Each atomic interaction in the framework is designed with explicit service dependencies. For example, authentication relies on services such as OAuth for secure token exchange, while communication may leverage WebRTC or RTP for real-time data transfer. Support services like haptic feedback engines or machine translation APIs further enhance these interactions. Consequently, the framework acts as an abstraction layer that binds user-facing functionality to underlying technical infrastructure in a standardized and re-usable way.

Bridging to use cases

The Framework for Virtual Worlds Interaction does not define specific use cases but provides the re-usable, modular interactions necessary to construct them. It serves as a toolkit that application developers and domain experts can combine and configure to meet contextual requirements. For example, a virtual government service centre may involve authentication, navigation, communication, consent management and feedback, each represented as atomic interactions defined by the framework.

This structured approach facilitates the systematic development of inclusive and interoperable use cases. By abstracting interactions from specific implementations, the framework promotes cross-domain compatibility and accelerates the development of standards-compliant applications. It also ensures that accessibility and inclusion are embedded at the interaction level, thus propagating these principles into every derived use case. Detailed use case configurations and domain-specific adaptations are addressed in Chapter 6.

4 Tools, terminologies and technologies in XR

eXtended Reality (XR) experiences rely on a diverse set of tools and technologies that bridge the digital and physical worlds, enabling natural and immersive interaction with XR systems. This section outlines such key input and output modalities – ranging from hand, face and eye tracking to bio-signals and haptic feedback. Understanding these technologies and their associated terminologies is essential to building XR systems that enhance accessibility, inclusivity and user engagement within virtual environments.

It is highly recommended to have disabled people on board when extending the use of assistive technologies. People with disabilities have more expertise and creative hacks in re-adaptation and modification of assistive technologies.

1) **Hand tracking** – A system to detect and interpret hand poses, gestures and finger movements of the user in XR.
2) **Hand-held controllers** – Physical devices used to navigate and interact with XR environments via buttons and motion tracking.
3) **Face tracking** – Captures facial expressions to enable realistic avatar animation and emotion recognition.
4) **Mouth and lip tracking** – Monitors movement around the lips and mouth for accurate speech representation and lip-sync in avatars.
    a. Tongue based interaction – Uses tongue motion for input, often in assistive or hands-free XR interfaces.
    b. Teeth based interaction – Detects teeth clenching or tapping as discrete control inputs in XR.

5) **Eye tracking** – Tracks eye movements to enable gaze-based interaction, foveated rendering and attention-based user interface elements in XR environments.

    a. Webcam-based Eye Tracking – Uses standard webcams to detect eye position and movements, often incorporating software techniques to compensate for head motion.

    b. Remote eye tracking – Operates without wearable equipment, using external cameras placed at a distance to monitor eye movements.

    c. Mobile eye tracking – Integrated into mobile or head-worn XR devices for tracking on the go, useful in untethered or field applications.

    d. Gaze tracking – Determines where the user is looking within a virtual environment, supporting features like gaze-based selection, user attention analysis and dynamic rendering.

6) **Feet tracking** – Captures foot position and motion for locomotion and full-body presence in XR.

7) **Full body tracking** – Captures and maps the entire body's movement to enable realistic avatar representation and physical interactions in XR. Full-body tracking systems may use a combination of visual, inertial, thermal and biosignal inputs to achieve accurate motion representation across varied environments and use cases.

    a. Optical tracking – Uses visual inputs such as cameras (Red, Green, Blue (RGB), infrared, or depth-sensing) to track body movement. Can be implemented in two modes:

        i. Inside-out tracking – Employs sensors and cameras mounted on the user (e.g., in a headset or body-worn device) to observe the surrounding environment and infer motion.

        ii. Outside-in tracking – Utilizes external cameras or sensors placed in the environment to track the user's movements.

    b. Inertial tracking – Leverages inertial measurement units (IMUs) such as accelerometers and gyroscopes, often worn on the body, to detect motion and orientation without relying on external cameras.

    c. Thermal tracking – Uses body heat signatures to infer full body movement and posture especially useful in low-light or non-visual environments.

8) **Voice based interactions** – Allows users to control and interact with XR using speech.

    a. Using context based input (e.g., Large language models (LLMs)) – Employs AI to interpret user intent from natural, contextual speech.

    b. Using basic command based input – Relies on predefined spoken commands for XR interaction.

9) **Bio signals-based interaction** – Uses physiological signals to interpret user states or enable control.

    a. Breath – Monitors respiratory patterns for interaction or emotional state detection.

        i. Air-flow through mouth – Detects breath changes via mouth for input or health monitoring.

        ii. Expansion of chest and abdomen – Tracks physical breathing motion using sensors or cameras.

    b. PPG – Photoplethysmography measures heart rate or blood volume via light-based sensors.

    c. EEG – Electroencephalography records brainwave activity for cognitive state monitoring or control.

   d. ECG – Electrocardiography tracks electrical activity of the heart for health or interaction cues.
   e. EMG – Electromyography senses muscle activity to detect gestures or movement intention.

10) **Brain-controlled interfaces** (BCIs) – Interfaces enabling direct control of XR using brain signals.
   a. Invasive Interfaces – BCIs involving implants interfacing directly with brain tissue.
   b. Non-invasive Interfaces – BCIs using external sensors to detect brain activity without surgery.
11) **Neuromuscular signals** – Uses subvocal muscle activity to silently communicate or interact.
12) **Existing assistive technology as an XR interface** – Integrates already-in-use assistive devices into XR for accessible interaction.
   a. White cane-based interaction – Extends traditional cane use into XR with spatial feedback and object detection.
   b. Wheelchair based interaction (e.g., chairable computing) – Uses wheelchair movement or controls as XR input.
   c. Sign Language based interaction – Recognizes sign language for communication and control in XR. (NOTE: Sign language recognition depends not only on hand movements but also on the facial and mouth expressions of the signer).
   d. People as affordances (a concept introduced by Arseli Dokumaci, 2023). For example, able-bodied people guiding people with disabilities in XR for movement and interaction. That is, the support of other individuals as assistive entities within XR experiences.
13) **Haptic feedback** – Provides physical sensations to simulate touch, force, or texture in XR.
   a. Mid-air haptics – Delivers touch sensations through air using ultrasound without contact.
   b. Electrical muscle stimulation – Uses electrical signals to stimulate muscles and convey touch or resistance.
   c. Haptic gloves – Wearable gloves that simulate tactile feedback and hand interaction in XR environments.
14) **Spatialized Sound as information** – Uses 3D audio cues to convey direction and context in virtual space.
   a. Contextualized information (e.g., directed awareness in VR) – Spatial audio guides user attention based on context or tasks.
      i. LLMs as directed feedback (e.g., move left towards the door) – AI-generated spatial voice cues provide personalized navigation or instructions.
   b. Non-contextualized information (e.g., alert system) – Delivers generic spatial alerts like warnings or notifications via sound.

5 Accessibility services

Accessibility services provide an alternative communication for users who cannot access or fully access audio or visual elements and for users who cannot fully understand standard language. Two key features need to be considered in virtual worlds:

1) **Positioning**: Access services need to be positioned in the virtual world. For instance, textual alternatives like subtitles (also called captions) need to be located on a certain area of the virtual sphere whereas audio alternatives like audio description also need to be positioned on a certain location, i.e., the audio needs to come from a certain area of the virtual sphere.
2) **Guiding mechanism**: A mechanism towards the focus of attention (e.g., speaker, virtual object, action) is needed to guide the user. For instance, if a person with hearing loss is navigating in the virtual world and someone is speaking in a location at their back, they would need a mechanism to guide them to the speaker. Examples of guiding mechanisms are arrows or radars. A light around an object could also be used as a guiding mechanism towards an object.

Additionally, XR environments shall ensure access through different input and output alternatives and facilitate user personalization to fulfil user needs.

5.1 Textual alternatives

Text alternatives for audio can take different forms. Subtitles are instrumental for users who do not understand the language (interlingual subtitles) and for those who do understand the language but cannot hear it (intralingual subtitles). In this case subtitles – sometimes also called subtitles for the deaf and hard-of-hearing (SDH) – provide an alternative for speech and non-speech information such as character and language identification, paralinguistic elements, sound, silence and music. Similarly, transcripts are written alternatives to spoken words, sometimes including non-speech information and generally in the same language of the audio. For example, a user may need a transcript when they participate in a virtual meeting on a virtual world.

5.2 Audio alternatives

There are different access services that provide audio alternatives to visual content, be they images or written text.

Audio description provides an auditory alternative to visual content: audio description describes the visuals and supports those who cannot access the images. Audio description can be applied to static content (e.g., a work of art) or dynamic content (e.g., a video). For example, users with sight loss enjoying an art experience in a virtual world would benefit from an audio description of the works of art. They would also need an easy way to navigate to the place where the virtual exhibition is taking place and an audio description of the environment.

Audio introductions provide an audio summary of the visual content before accessing it. Audio introductions complement audio descriptions. For example, a user may listen to an audio introduction describing the main features of a virtual world where a gaming experience will take place before entering this virtual world.

Visual content can sometimes be written text. For instance, content in a language A may be subtitled in a language B for those who do not speak the language. However, some users may have difficulties seeing or reading the subtitles and on-screen text. Audio subtitles provide an audio version of subtitles and on-screen text for those who do not understand the source language and cannot read these visual textual elements. For example, a virtual world may feature an avatar speaking in Catalan with automatic subtitles in English. An English-speaking person with sight loss or with reading difficulties would not be able to access this content unless audio subtitles (also called spoken subtitles) are offered.

5.3 Visual gestural alternatives

Sign language interpreting provides a visual gestural alternative to oral languages, benefitting users who cannot access the audio and understand that sign language. Similarly, oral language interpreting may provide a spoken version of a sign language or another oral language.

For instance, spoken content in English in a virtual world may be signed into American Sign Language. Similarly, a Catalan sign language speaker and a Catalan oral speaker may communicate through interpreting services.

5.4 Haptic alternatives

Haptic alternatives may facilitate and enhance communication with users through vibrotactile features or by a change of temperature. For instance, haptic feedback may offer cues to a blind user so that navigation in the virtual world is easier. Deaf-blind users may use a Braille display to access visual information.

5.5 Easy-to-understand languages

Easy-to-understand languages encompass different simplified language varieties that enhance comprehensibility, ranging from Easy Language (or Easy-to-Read or Easy Read) to Plain Language. Easy Language is the most simplified modality and is especially addressed to persons with comprehension difficulties for any reason. Plain Language is less simplified and is addressed to a broader user group, not necessarily with comprehension difficulties.

In a virtual world, it would be good practice to offer instructions in Plain Language, providing also Easy Language alternatives. Similarly, it would be necessary to offer users an easy way to go to a safe place or allow them to customize the environment in the virtual world in case they do not understand or feel overwhelmed.

For instance, a user may look for a more relaxed gaming experience in the virtual world, with easy navigation alternatives, a slower pace, Easy Language content and a shortcut that takes them to a safe space or adjusts the light or the noise.

Easy-to-understand languages can also be applied to access services such as audio description, subtitling or sign language interpreting to create hybrid easy-to-understand access services.

5.6 Final considerations

The previous section describes some accessibility services providing alternatives, but the list is not exhaustive. New possibilities and new user needs may emerge as virtual worlds evolve. For instance, the increasing use of AI can act as an accessibility service enabler in relation to language and visual recognition and generation. Special care should be taken as new human-computer interactions (HCI) emerge.

Access services also need to consider questions related to ethics and avoiding bias, as well as facilitate the sense of presence and co-presence in virtual environments.

Access services can be seen as an alternative for users who cannot access one mode and as an enhancement for those who can access it. For example, an image of a ball bouncing can be translated into vibrotactile feedback. This haptic experience can also enhance the experience of persons without sight loss.

Moreover, it should be noted that access to the VR equipment is not possible for all users, hence the need to consider how access will be facilitated.

Table 1 summarizes the accessibility services presented in this section.

Table 1: Access services linked to communication needs

Communication need	Type of alternative	Access service
Access to audio content	Textual	Caption/subtitles Transcripts
Access to visual/textual content	Audio	Audio description Audio introduction Audio subtitling

Table 1: Access services linked to communication needs (continued)

Communication need	Type of alternative	Access service
Access to audio content	Visual-gestural	Sign language interpreting
Access to visual content	Haptic	Haptic (vibrotactile) Tactile (Braille)
Access to complex content/ audio/visual/textual)	Easy-to-understand	Easy Language Plain Language
Access to audio content	Textual	Caption/subtitles Transcripts

6 Use cases

As immersive technologies such as Virtual Reality (VR, Augmented Reality (AR), Mixed Reality (MR) and Extended Reality (XR become increasingly integrated into education, workforce training and cultural experiences, so the importance of accessibility in these environments becomes ever more pressing. While immersive systems promise unprecedented levels of engagement and interactivity, they also risk excluding individuals with disabilities unless inclusive design is considered from the outset. This section presents a collection of case studies showcasing how various organizations, institutions and communities have tackled accessibility challenges in immersive environments. Each case demonstrates the critical importance of co-design, the innovative application of assistive technologies and the emerging role of artificial intelligence in creating equitable digital experiences.

6.1 Accessible learning platforms: students with disabilities in immersive environments

6.1.1 Overview

As immersive technologies become increasingly central to digital education, it is vital to consider how these platforms can be made accessible to all learners. For students with disabilities, the shift to immersive environments introduces new challenges and opportunities for inclusion. This case study explores the potential for designing accessible immersive learning platforms, discussing possible strategies, anticipated barriers and the importance of embedding inclusive design from the outset.

6.1.2 Background and context

While significant progress has been made in the accessibility of websites and traditional learning management systems (LMS), immersive environments remain relatively underexplored in this regard. Unlike conventional platforms, immersive learning experiences often rely heavily on visual, spatial and motion-based interaction, which can inadvertently exclude students with disabilities. For example, learners with visual impairments may struggle to navigate virtual spaces without detailed auditory feedback or spatial sound cues. Students with hearing loss might find themselves unable to fully engage with audio-based simulations that lack captioning or visual indicators. Similarly, individuals with limited mobility may find standard hand-held controllers difficult or impossible to use and those with neurodivergent cognitive styles may find complex 3D environments disorienting or overwhelming.

Given these concerns, there is a growing imperative to ensure that immersive education evolves in an inclusive direction. While accessibility in digital education has typically focused on screen readers, alternative text and captioning, immersive platforms introduce a need for a broader, multisensory understanding of accessibility. Addressing these needs proactively could help ensure that immersive education empowers all students rather than creating new exclusions.

6.1.3 Objectives

A future initiative to improve accessibility in immersive learning environments would likely aim to ensure that students with a wide range of disabilities can engage meaningfully and independently with immersive content. This would involve developing inclusive interaction models that accommodate a variety of physical, sensory and cognitive abilities. It would also necessitate support for alternative input methods such as voice commands, gaze tracking, adaptive switches and keyboard navigation, along with alternative outputs including audio descriptions, haptic feedback and real-time captioning. Beyond the technology itself, such an initiative would need to address pedagogical practice by equipping academic staff with training and tools to create accessible immersive content from the outset. As these environments become more sophisticated, there is also an opportunity to explore the use of AI to personalize learning experiences and adapt environments in real time to meet individual accessibility needs.

6.1.4 Implementation

To implement this within an institution, a possible approach would begin with a comprehensive audit of current immersive learning content to identify where accessibility gaps exist. From there, universities could establish cross-functional teams comprising learning designers, accessibility specialists, immersive developers and students with lived experience of disability. These teams would be tasked with developing a framework for accessible immersive learning grounded in emerging standards such as those promoted by the XR Access Initiative, while also drawing on principles from existing accessibility guidelines adapted to the unique challenges of 3D environments.

In practice, immersive learning scenarios might be designed from the outset to include features such as real-time audio description of visual elements, adjustable user interfaces with high-contrast and simplified layouts, closed captions embedded directly into VR and AR scenes and compatibility with assistive technologies such as eye trackers and switch controls. Input devices could be mapped flexibly to accommodate different motor abilities and environmental complexity could be scaled or toggled to suit different cognitive needs.

Crucially, this implementation would involve iterative prototyping and user testing with disabled students. Their feedback would not only help to validate whether solutions work but also guide the creative process to ensure the learning environments are genuinely inclusive and intuitive. Training and capacity building would be needed across the academic community, helping educators understand how to integrate accessibility into immersive learning design using platforms such as Unity or WebXR.

6.1.5 Challenges and barriers

Several challenges are likely to emerge when attempting to embed accessibility into immersive learning. A primary difficulty is the lack of well-established accessibility standards for VR, AR and MR platforms, making it harder to evaluate and benchmark progress. Technical limitations are also a concern; many immersive tools are not yet designed with assistive technologies in mind and features like real-time audio description or AI-based content adaptation are still in their infancy. Furthermore, the level of expertise required to build immersive content is already high, and adding accessibility considerations may be perceived by some faculties as an additional burden, particularly in the absence of clear guidance or support.

Cost presents another significant barrier. Developing accessible immersive content often requires additional resources – such as haptic hardware, adaptive input devices, or captioning services – that may not be readily available within university budgets. Moreover, legacy immersive materials are often not designed with accessibility in mind, meaning that retrofitting them can be complex and resource-intensive.

Cultural resistance may also play a role. Immersive learning is frequently positioned as an experimental or cutting-edge domain and accessibility is sometimes viewed as secondary to innovation. Addressing this mind-set will be key to ensuring accessibility is embedded from the start rather than treated as an afterthought.

6.1.6 Outcomes and impact

If these challenges can be addressed, the potential benefits are substantial. Accessible immersive learning environments could enable students with disabilities to engage in experiences that were previously inaccessible such as virtual science laboratories, field trips or simulated professional scenarios. These environments could be configured to support a wide range of needs, allowing learners to control the pace and complexity of interactions, to switch between sensory modalities or to receive targeted support through personalized AI-driven adaptations.

Such improvements would not only benefit disabled students but also enhance the learning experience for all. Features like captions, simplified navigation, and customisable controls have been shown to support broader groups of learners, including non-native speakers, students with anxiety and those with different learning preferences. In time, a university that embraces this inclusive approach could emerge as a leader in accessible immersive education, so contributing to sector-wide standards and demonstrating the educational value of inclusive design.

6.1.7 Lessons to consider

Accessibility must be treated as a foundational element of immersive content design, not as something added once the technology is complete. Co-design with students who have disabilities is essential for creating learning environments that are not only technically accessible but also pedagogically meaningful. The collaboration between technical developers and educational practitioners must be close and continuous, as each group brings critical insight into how users interact with content. Furthermore, while adaptive technologies and AI hold promise, they are not a substitute for human-centred design processes. Institutional commitment – supported by policy, leadership, funding and training – will be necessary to maintain momentum and ensure that accessibility becomes a core expectation in immersive teaching and learning.

6.1.8 Case study conclusion

This case study highlights the importance and the feasibility of developing accessible immersive learning platforms in higher education. As universities continue to explore the possibilities of VR, AR and MR in teaching, it is imperative that accessibility is not left behind. While significant technical and cultural challenges remain, there is a clear pathway toward inclusion, guided by principles of universal design, co-creation and institutional responsibility.

By acting now, institutions can shape the future of immersive learning as an inclusive, empowering space for all students. Further research, collaborative development and sustained commitment will be essential to ensure that immersive learning realises its potential not just as a technological innovation, but as a force for educational equity.

6.2 Immersive training: Bank of America's integration of virtual reality and AI in workforce development

6.2.1 The challenge

Bank of America[41] sought to equip its extensive workforce with the skills and confidence necessary to handle complex customer interactions and high-pressure situations effectively. Traditional training methods lacked the realism and engagement required to prepare employees adequately for these challenges.

6.2.2 The opportunity

Advancements in VR and AI technologies presented an opportunity to revolutionize employee training by providing immersive, realistic simulations. These tools could offer employees hands-on experience in a controlled environment, enhancing their ability to manage real-world scenarios with competence and empathy. Equally important, such platforms opened new possibilities for accessibility, including multilingual interfaces and voice-guided simulations tailored to users' language preferences.

6.2.3 The action

In 2021, Bank of America became the first financial services firm to launch a VR training programme across nearly 4 300 financial centres nationwide, reaching approximately 50 000 employees. The VR modules allowed employees to practice a range of skills, including strengthening client relationships, navigating difficult conversations and responding with empathy. The programme included 20 different simulations designed to replicate real-world scenarios such as handling customer complaints or responding to security threats. Building on this foundation, by 2023, the bank expanded its training initiatives to incorporate AI and metaverse technologies. New hires participated in immersive training programmes where they engaged with AI-powered bots simulating client interactions, providing real-time feedback and coaching. These simulations acted as practice repetitions, enabling even new employees to gain experience and confidence quickly. To promote inclusivity, the bank also introduced support for multiple languages and embedded accessibility features such as closed captions and screen-reader compatibility to ensure the training was usable by employees with diverse needs and abilities.

6.2.4 The results

The implementation of VR and AI training programmes yielded significant benefits:

- Enhanced Confidence and Performance: Following the initial VR training pilot, 97 per cent of participants reported feeling more comfortable performing their tasks.
- Improved skill retention: Employees demonstrated higher retention rates of training material compared with traditional methods, leading to better preparedness in real-world situations.
- Scalability: The programme successfully scaled to train more than 200 000 employees globally, providing consistent and effective training across the organization.

6.2.5 Lessons learned

Bank of America's experience underscores the effectiveness of immersive technologies in employee training. Key takeaways include the importance of starting with pilot programmes to assess impact, the value of realistic simulations in building employee confidence and the necessity of continuous innovation to meet evolving training needs.

Additionally, incorporating accessibility features from the outset expanded the reach and impact of the training programmes, reinforcing the importance of inclusive design in digital transformation efforts.

6.2.6 Looking ahead

The bank continues to explore the use of immersive technologies for various applications, including virtual onboarding and team building. By leveraging the metaverse, Bank of America aims to enhance employee engagement, improve customer service and maintain its position as a leader in innovative workforce development strategies. Future enhancements will further prioritize accessible design such as adaptive learning paths and localized training environments, to ensure all employees, regardless of language or ability, can fully benefit from the experience.

6.3 Accessible cultural heritage visits in virtual worlds

6.3.1 Overview

Virtual worlds can allow tourists to visit a city and its cultural heritage without leaving home. This opens up many possibilities to people who, for multiple reasons (e.g., disability, environmental impact), cannot, or do not want to, move to a different location but still want to enjoy cultural sites. There is an opportunity to make these virtual cultural visits accessible for all, avoiding the barriers still frequently found in physical locations. This use case study addresses how an accessible cultural heritage visit was planned and developed from the cultural heritage site perspective in order to enhance tourism for all in virtual worlds.

6.3.2 Background and context

Cultural heritage sites have long struggled to make their spaces accessible to all visitors, including people with disabilities and the elderly. To achieve this aim, actions to remove physical barriers and incorporate accessibility services have been taken by some institutions. When moving to a virtual world, some challenges remain but new ones emerge. Critical steps such as planning the visit, accessing the virtual world, navigating and interacting with the cultural site/ environment and providing feedback after the virtual visit were considered in this use case to cater for a truly accessible experience.

6.3.3 Objectives

The primary goal of this initiative was to ensure that anyone, including persons with disabilities, could enjoy a visit to a heritage site in a virtual world. Specifically, the case aimed to consider critical accessibility steps in the four main stages in any cultural visit (planning, accessing, navigating, interacting and providing feedback) and integrate accessibility services by design.

6.3.4 Implementation

The first step was identifying the critical aspects and user needs for each of the four main phases in any cultural visit, always in collaboration with the end users.

1) When dealing with planning: Alternative communication services were requested by users who also highlighted the need for multilingual content. Accessible information about the virtual cultural heritage site and easy-to-understand instructions on how to access the virtual experience were provided, responding to user needs, and following existing standards and guidelines such as EN 301549, WCAG 2.2. and XR Accessibility User Requirements (XAUR).
2) When dealing with accessing the virtual experience: Users required the possibility to access the experience through several types of device, with alternative options to authenticate themselves. They highlighted the need for hardware components to be interoperable with assistive devices. The creation of an avatar was the subject of some interesting discussions, as some users wanted to be represented with their guide dog or a cane, whereas others wanted to explore new representations, highlighting the need to allow for self-representation or an easy control to turn on/off, or switch between specific features.
3) When dealing with navigation and virtual site/environment interaction: Users pointed at the need to offer different navigational and guidance options. Similarly, users expressed the need to input and receive responses by alternative options, depending on their preferences. The challenge of multilingualism was also put forward, as users may need to communicate and receive information about the environment and objects not only through different channels but also through different languages. Interactions between users was considered relevant but outside the scope of this use case.
4) When leaving the virtual experience: Users suggested the implementation of accessible feedback mechanisms in which areas for improvement could be identified.

The second step was the actual implementation of the virtual world, with associated accessibility services and accommodation/adjustment options, as per the user needs and requirements.

The third step was the testing with diverse users of the virtual world experience.

After successfully concluding this third step, the virtual site was launched, with its associated information announced in an accessible manner.

6.3.5 Challenges and barriers

The implementation process was not without its difficulties and funding concerns were expressed by some. Technical implementation was not without difficulties and interoperability with assistive technologies was also a main barrier as immersive environments can be developed through different technologies and tools. The fact that users were involved from the beginning, rather than adding accessibility as an after-thought, eased the process.

6.3.6 Outcomes and impact

The initiative reached new audiences and allowed some persons with difficulties reaching physical sites to enjoy cultural heritage (for instance, elderly). The virtual site was also used by primary and secondary teachers in their teaching. Feedback from visitors was positive. Feedback from the cultural heritage site professionals was also positive, as they acknowledged increased awareness about diverse user needs.

6.3.7 Lessons learned

Accessibility should not be considered an afterthought but rather as an integral part of the design process. There is a need for simplification in virtual worlds to avoid information overload. Questions related to movement of the user (locomotion) within the virtual worlds may lead to cybersickness. There is a need to ensure proper representation and integrate proper feedback mechanisms throughout the visit. Accessibility services to allow user access, navigation and interaction in virtual worlds should be ensured.

6.3.8 Case study conclusion

This case study demonstrates the requirement to cater for the accessibility of all users in virtual environments regardless of their needs.

6.4 e-Extended Reality training for automation

European industry's investment in advanced technological, resilience, entrepreneurial and ecological skills is vital for the inclusive green and digital transitions. Advanced Manufacturing Technologies (AMT) and the Deep Tech approach, converging digital and physical realms, are crucial. Industrial operators need new emotional and relational skills for interacting with intelligent machines, while training designers require a holistic, multidisciplinary approach. Extended Reality and the Industrial Metaverse (XR/IM) are key to innovating production and fostering the skilled, resilient and creative workforce that is essential for successful industrial transitions and competitiveness. Designing industrial virtual spaces which are accessible for all will work towards a culture of integration and avoiding exclusion, in particular with relation to physical accessibility, by building virtual spaces that are accessible for people with disabilities, without barriers like ramps and inappropriate signage and digital accessibility, by overcoming software usability issues, often experienced by those workers with visual, auditory, or motor impairments.

6.4.1 The challenge

Industry training for automation, specifically programming a metal bending robotic cell, utilizes a practical approach combining CNC (Computerized Numerical Control) and collaborative robot programming with hands-on exercises. This experience-based method accelerates learning compared to traditional methods.

6.4.2 The opportunity

The case envisions integrating XR for a continuous learning flow across physical, virtual reality and AR. Trainees learn on-site with AR guidance, then immerse in VR for risk-free practice and problem solving, finally applying skills in the real world and collaborating in a digital workspace.

6.4.3 The action

A pilot experiment at a modular robotic cell with a metal sheet bending machine tested XR technology for industrial training. Participants with no prior machine knowledge used smartphones and an AR app to document robotic cell programming tasks via videos and photos, replacing traditional notes. Hands-free recording glasses were also tested for automatic video uploads. Subsequently, a VR application for managing training notes was evaluated for design and usability feedback. Later, operators utilized the AR app to access training notes and asynchronously request remote assistance from trainers. Finally, the AR mobile app was presented to stakeholders to explore its potential integration into their service offerings and business strategy.

6.4.4 The results

The case study revealed accessibility, usability and effectiveness strengths in AR and VR for knowledge note manipulation, leading the trainer to emphasize the smoothness of the transitions among the realities. The focus on XR technology accessibility brings developers and designers attention to the enabling requirement for XR: Internet connection availability that makes the web applications and services accessible and effective. Despite varying initial experience, trainers and trainees quickly learned to respectively operate the XR authoring application and the end-user bottom-up note taking on robotic bending cells. In particular, augmented reality significantly improved comprehension by linking notes to machine parts and the whole ecosystem facilitated organized information tracking and sharing, creating a valuable knowledge base for the organization.

6.4.5 Lessons learned

The case study highlights XR technology's potential to empower production workers by enabling them to perform tasks requiring advanced skills such as complex machinery setup and diagnostics, where current capabilities for autonomous intervention are limited in SMEs. Developing effective XR training necessitates addressing the diverse challenges present within these smaller companies. Consequently, creating a framework encompassing appropriate operator skills and supportive technologies is crucial for equipping younger generations with the competencies needed to work effectively, safely and with resilience in advanced manufacturing environments.

6.4.6 Looking ahead

Building on the preference for AR and the need for a unified training solution, the next research step should focus on integrating AI for personalized automation training. This involves developing an AI-powered system that assesses a user's existing knowledge and expertise with industrial machines. The system would then dynamically adapt the training content, pace and difficulty, delivered through an AR interface. This personalized approach aims to optimize learning outcomes, address individual skill gaps more effectively and further empower production workers in SMEs.

6.5 Enabling mouth-based interaction in XR

Following the rise of affordable consumer-focused XR technologies disabled artistic communities have started exploring XR platforms for creative expression and immersive experiences. However, the reliance on head movements for 360-degree point of view and rapid hand movements to interact with these devices still excludes people with motor disabilities such as muscular dystrophy or spinal cord injuries.

6.5.1 The challenge

Eric Desrosiers and Christian Bayerlein are two disabled artists living in Montreal, Canada and Koblenz, Germany respectively. Eric Desrosiers is a painter in his mid-50s who has muscular dystrophy, a degenerative neuromuscular disease and Christian Bayerlein is an artist, technologist and activist who has Spinal Muscular Atrophy (SMA). Eric and Christian, with their disabilities, have almost no movements in the upper and lower limbs of their body. However, both artists regularly use various technologies using their mouth: mouth-operated wheelchairs to navigate and mouth-controlled assistive joysticks to control robotic arms and integrate such devices into their artistic practices. However, the reliance on hands and finger gestures to access the current generation of XR joysticks/controllers, restrict them to use the immersive potential of technologies such as Virtual/Mixed Reality in their art. Can the citiverse deliver this promise of inclusivity and provide artists like Christian and Eric to participate fully? This is the challenge addressed by this case study.

6.5.2 The opportunity

The case envisions integrating a set of mouth-gestures using AI and camera-based interfaces in XR. Such a feature would henceforth enable artists such as Eric and Christian to use their mouth to move in virtual spaces and interact with virtual elements solely using their cheeks, tongue and lip movements. This is extremely beneficial for them for two reasons: 1) they cannot rotate or move their heads to have a 360-degree point of view in VR; 2) they already use their mouth to operate various assistive technologies in their everyday life. Hence, beyond contributing an accessibility feature in XR for artists who rely on their mouth for their everyday interaction with technologies, this case study also opens opportunity to discover novel interaction modalities in XR (e.g., mouth) which currently are predominantly reliant on hand and eye movements.

6.5.3 The action

To detect mouth gestures, firstly, open source facemesh landmark detection algorithms were used to track the face; in particular, the contours of the mouth. Custom statistical algorithms were developed to classify certain gestures such as opening of the mouth, puffing the cheeks and moving lips laterally. These gestures were later mapped to specific virtual camera movements in VR scene to move forward (mouth open), rotate (cheek puff) and left/right (lateral movement of the lips). However, due to the inability of the available facemesh landmark detection to recognize occluded faces (e.g., when the head-mounted display covers the eyes and nose), facemasks were pasted on the headsets to make the algorithms work. Lastly, an existing mouth-tracking device was re-modified to detect six mouth gestures to move in VR and interact with virtual elements in mixed reality (e.g., switching on/off the "passthrough" feature of mixed reality headsets for Eric and Christian). Furthermore, these gestures were integrated into a co-developed artwork (with non-disabled and disabled engineers and artists), titled *Crip Sensorama*,[42] that puts the accessible developed interfaces in the service of such artists to showcase their art.

6.5.4 The results

The case study illustrates that current XR technologies remain largely inaccessible to individuals with severe motor disabilities. This study demonstrates a pathway toward more inclusive design by collaborating closely with two disabled artists to modify and re-engineer existing camera-based interfaces and AI algorithms to develop sets of mouth gestures for the artists, enabling them to navigate, move and interact in XR solely using their mouth movements. Furthermore, leading to the creation of XR artwork *Crip Sensorama* exhibited in 2024, enabling the (non-disabled and disabled) public to not only adapt unconventional interaction modalities in XR (e.g., mouth gestures) but also to experience disabled art and practices through the same.

6.5.5 Lessons learned

The design process enabled the engineers on the team in understanding accessibility in XR and the Citiverse in three ways:

1) Prioritizing lived experience(s) as core design input in accessibility technology: Instead of using common-modalities such as voice commands, eye-gaze or the micromovements of fingers as an accessibility feature, Eric and Christian have a heavy reliance on physical mouth movements in their daily lives, which enabled the engineers to focus on mouth gestures, leading to the development of an unconventional interaction modality in XR.
2) Adapting Prototypes by Embracing Users Creativity: Existing machine-learning algorithms were creatively re-purposed; for example, the face-mesh landmark detection system (e.g., pasting facemasks), which typically fails when facial features are occluded. This was made possible by acknowledging artistic inputs from the artists, hence discovering imperfect yet functional hacks such as using facemasks on the headsets to track the face, and hence the mouth.
3) Using Assistive Interfaces in Public for Disability Empowerment and Inclusion: The developed camera-based interfaces were used in publicly exhibited art works which enabled the visitors to appreciate novel modalities of interaction in XR but also enhance their understanding of disability art and accessibility.

6.5.6 Looking ahead

To enhance the understanding of mouth-based navigation and interaction in XR, more complex gestures are still left to be investigated. As a running example, while extending the tongue out could continue to move the user forward, opening the mouth midway could facilitate jumping while in motion within VR.

6.6 e-Extended Reality boxing: An accessible VR platform for blind users

Virtual Reality offers immersive experiences through advanced interaction mechanisms and rich sensory stimuli but is often inaccessible to blind people due to its over-reliance on visual feedback. While prior work has investigated specific aspects of VR accessibility, there is little knowledge on how to design full, feature-rich VR experiences accessible to blind people. This case study presents a pioneering approach to creating an accessible VR boxing experience for blind users, demonstrating how virtual environments can be designed to be inclusive and engaging.

6.6.1 The challenge

Mainstream VR applications rely heavily on visual feedback, making them largely inaccessible to blind users. Creating accessible VR experiences demands overcoming significant barriers, including providing spatial awareness, enabling natural movement and delivering rich feedback without relying on visual cues. Boxing, with its physical nature and dynamic interactions, presents a particularly complex challenge for accessibility, while offering significant potential for an engaging experience.

6.6.2 The opportunity

The physical nature of boxing makes it particularly well-suited for VR adaptation for blind users. The sport naturally leverages VR's affordances through user movement (locomotion and rotation) and physical interactions (punching and defending). This creates opportunities for multimodal feedback (audio and haptic) and dynamic interactions that do not primarily rely on vision. By developing an accessible VR boxing experience, we can demonstrate design principles applicable to a wide range of VR applications.

6.6.3 The action

A participatory design approach with an ex-professional boxer who became blind informed the development of a three-mode VR boxing experience.[43] The platform uses spatialized audio feedback, haptic responses through controller vibrations and movement-based interactions to create an accessible experience. Three progressive modes were implemented:

1) **Heavy bag training**: An entry-level experience where users practice punches on a stationary bag, receiving audio and haptic feedback.
2) **Coach training**: A dynamic training experience with a virtual coach providing punch instructions and an opponent that moves around the user.
3) **Combat**: The most complex mode featuring an opponent that moves freely around the ring, with audience sounds, the opponent's coach and more advanced movement requirements.

6.6.4 The results

A user study with 15 blind participants revealed high levels of engagement and enjoyment across all three modes (ratings of 6.4/7, 6.5/7 and 6.6/7 respectively). Participants particularly valued natural movement through real walking and turning, alongside the physical actions of punching and defending. Also mentioned was the use of the sound of the opponent's breathing and steps to convey their localization. The progressive complexity across modes allowed for structured skill development, with the virtual coach proving crucial for guidance and immersion. Key elements identified as most important included audio feedback of punches, the coach's instructions on positioning and audio cues indicating opponent movement.

6.6.5 Lessons learned

The project demonstrated several key principles for accessible VR design: (1) natural movement is essential for immersion; (2) rich auditory feedback can effectively replace visual cues; (3) well-timed guidance that balances assistance with user agency is critical; (4) structured progression enables users to build skills gradually; and (5) movement-based interactions combined with audio feedback create genuinely immersive experiences regardless of visual ability.

6.6.6 Looking ahead

Future development will focus on enhancing the spatial audio capabilities, adding more personalized coaching feedback and developing additional training scenarios. The identified design principles will be applied to other VR experiences beyond boxing. Additionally, the platform shows potential as a rehabilitation and training tool for spatial awareness and coordination, presenting opportunities for therapeutic applications. By demonstrating that rich, complex VR experiences can be made accessible, this project lays groundwork for more inclusive development across the VR landscape.

6.7 Summary

Table 2: Summary of Use Cases

Case	Title	Category	Primary focus	Target users	Accessibility features	Technology used	Co-design involvement	Scalability	Innovation highlight
6.1	Accessible Learning Platforms	Inclusive Education Technology	Embedding accessibility in immersive higher education	Students with disabilities	Voice/gaze control, audio descriptions, captions, simplified UI, haptics	VR, AR, WebXR, AI	Yes – students with lived experience	High (institution-wide potential)	Multisensory accessibility model & AI personalization
6.2	Bank of America VR/AI Training	Corporate Training and Workforce Accessibility	Enhancing staff performance using immersive training	General workforce (implicit inclusion)	Realism, repetition, consistency (no explicit disability features)	VR, AI, Metaverse	Limited (top-down design)	Very High (200k+ users)	Large-scale corporate deployment & AI coaching
6.3	Accessible Cultural Heritage Visits	Cultural Inclusion and Virtual Tourism	Making virtual tourism inclusive from planning to feedback	Tourists with disabilities, elderly, educators	Multilingual content, device compatibility, avatar customization, guided navigation	Virtual World Platform, WCAG, EN 301549	Yes – early and iterative user input	Medium (specific to site/platform)	Full-journey inclusive virtual tourism model
6.4	e-XR Training for Automation	Industrial XR Training and Knowledge Management	Accelerating technical training for industrial operators	Industrial trainees and SMEs	Interoperability, video-linked notes, AR assistance (accessibility emerging)	AR, VR, mobile apps, CNC machines	Limited – pending accessibility integration	Medium (pilot in industry context)	Use of AR for bottom-up industrial training documentation

(continued)

Case	Title	Category	Primary focus	Target users	Accessibility features	Technology used	Co-design involvement	Scalability	Innovation highlight
6.5	Mouth-Based XR Interaction	Assistive Interaction Design for Artists	Empowering artists with motor impairments in XR	Artists with severe motor disabilities	AI-based gesture detection, mouth gestures, headset mods	AI, face-mesh tracking, custom XR controllers	Extensive – lived experience central to design	Low-Medium (specialist application)	Novel mouth-based interaction paradigm for XR
6.6	e-Extended Reality Boxing: An Accessible VR Platform for Blind Users	Inclusive Entertainment and Gaming	Giving people with visual impairment access to e-sports	People with visual impairments	Rich VR interaction based on audio and haptic features	VR, Unity	Yes, from early requirements to multiple prototype iterations	High (accessible techniques implemented using consumer grade technology)	Rich non-visual interactive experience in what is usually a visual medium

6.8 Conclusion

Together, these case studies illustrate a cross-sector movement toward more inclusive immersive environments. Whether addressing educational barriers, redefining employee training, or empowering artistic expression, the common thread is a commitment to designing XR technologies that accommodate diverse users from the start. Co-creation with individuals with lived experience of disability consistently emerges as a best practice, enabling innovations that benefit all users, not just those with access needs.

While AI and immersive tools hold great promise, their deployment must be grounded in human-centred design and supported by institutional policies and resources. Accessibility should not be retrofitted; it must be an integral part of development. These case studies show that with foresight, creativity and commitment, immersive environments can become powerful tools for equity, inclusion and empowerment.

- **6.1 and 6.3** take the most **holistic view of accessibility**, covering the entire user journey and engaging disabled users from the outset.
- **6.2** focuses on **scalable immersive training** and while not explicitly accessible, it demonstrates a framework that could integrate accessibility features.
- **6.4** shows **promise in industrial training**, though its accessibility dimension is under development.
- **6.5 and 6.6** stand out for their **radical inclusivity and innovation**, introducing entirely new interaction modalities for entertainment driven by disabled artists and athletes.

Further research, collaborative development and sustained commitment will be essential to ensure that immersive technology realises its full potential, not just as a technological advancement but also as a force for educational, professional and creative equity.

7 Standardization

This section examines the development trends of international accessibility standards led by key organizations such as ITU, ISO/IEC JTC1 and W3C. Standardization in ICT for digital inclusiveness has primarily focused on accessibility for people with disabilities. Since the late 1990s, various international standards have been developed to enhance the accessibility of ICT products and services, by ITU-T Study Group 21 (SG21, formerly SG16) and ITU-T Study Group 20 (SG20) of the International Telecommunication Union, as well as by Subcommittee 35 (SC35) under the Joint Technical Committee 1 (JTC1) of ISO/IEC and World Wide Web Consortium (W3C).

ITU has developed international standards (called ITU-T Recommendations) on accessibility to telecommunication networks and services, while ISO/IEC JTC1 has developed accessibility standards across a wide range of IT domains, including hardware, software and the web, providing a reference framework for national adoption.

Although the World Wide Web Consortium (W3C) is not formally part of international standardisation development organisations (SDOs), it plays a crucial role in accessibility, especially web accessibility, which is the starting point and the most mature area of digital accessibility discourse.

W3C's Web Content Accessibility Guidelines (WCAG) have become the de facto international standard, recognized as ISO/IEC 40500 and adopted by numerous countries as national standards, forming a shared global framework for accessibility.

7.1 ITU

As a UN specialized agency for digital technologies, ITU has published standards for ICT accessibility since the mid-2000s. Within ITU's Telecommunication Standardization Sector (ITU-T), various study groups have developed standards for accessible communication services targeting persons with disabilities and older persons. In 2007, ITU-T approved its first comprehensive accessibility guideline: Recommendation ITU-T F.790, "Telecommunications accessibility guidelines for older persons and persons with disabilities"[44]. This document guides designing accessible communication devices, software and services, serving as a global reference for telecom operators and manufacturers. It includes a wide range of requirements – from the physical design of telephone terminals to software user interfaces and service procedures – forming the foundational framework for subsequent accessibility standards.

ITU has since expanded its scope to cover media accessibility and terminology. In 2015, ITU approved Recommendation ITU-T F.791 "Accessibility terms and definitions"[45], which defines standardized terms related to accessibility, disability and assistive technologies, aiming to support consistency among national standards developers. In the broadcasting and multimedia domain, Recommendation ITU-T H.702 (2015) "Accessibility profiles for IPTV systems"[46] defines accessibility profiles for IPTV systems, including support for subtitles, sign language and audio description; the 2020 revision updated technical requirements for delivering accessible content on modern IPTV platforms. Additional specifications include annexes to the H-series defining testing frameworks and FSTP-AM (2015), a guideline on audiovisual conferencing accessibility.

Several standards targeting audiovisual impairments were issued in 2018. Recommendation ITU-T F.921 "Audio-based indoor and outdoor network navigation system for persons with vision impairment"[47] defines concepts and requirements for audio navigation systems for visually impaired individuals, applicable to indoor and outdoor environments. Recommendation ITU-T F.930 "Multimedia telecommunication relay services"[48], on the other hand, specifies functional and procedural standards for multimedia telecommunication relay services such as remote sign language interpretation for individuals with hearing or speech disabilities. Other reports such as FSTP-ACC-RemPart (2015) on remote participation accessibility, have gained relevance during the pandemic and the era of increased virtual engagement.

ITU is also actively developing accessibility standards in emerging fields like smart cities, the Internet of Things (IoT) and artificial intelligence. Recommendation ITU-T Y.4204 "Accessibility requirements for the Internet of things applications and services"[49], developed by ITU-T Study Group 20 (IoT, digital twins and Smart Sustainable Cities and Communities)[50], outlines accessibility requirements for IoT-based services for persons with disabilities and the elderly. It presents various use cases – such as smart home devices, wearables and environmental control systems – and categorizes design considerations by types of disabilities (visual, hearing, motor and cognitive). In 2023, this was expanded into Recommendation ITU-T Y.4219 "Accessibility requirements for user interface of smart applications supporting IoT"[51], which adds further guidance to enhance accessibility across broader IoT and smart city infrastructures. To ensure accessibility is integrated into all ICT standards, ITU has also adopted ISO/IEC Guide 71:2014 as Supplement ITU-T H. Suppl.17, which recommends incorporating accessibility considerations into all stages of standards development in cooperation with ISO and IEC.

7.2 ISO/IEC JTC 1

The ISO/IEC JTC 1 (Joint Technical Committee 1 on Information Technology) formulates international standards encompassing all domains of Information and Communication Technology (ICT) and has instituted accessibility standards in diverse sectors, including software, hardware, the web, education and assistive technologies through its dedicated working groups. Among the early standards established in the late 2000s were guidelines for software accessibility and documents delineating the requirements of users with disabilities. For instance, ISO 9241-171:2008 serves as a standard that provides comprehensive guidelines to guarantee that software interfaces can be utilized efficiently and satisfactorily by persons with disabilities. It incorporates extensive recommendations such as providing alternative text for user interface elements, supporting keyboard navigation and maintaining adequate contrast ratios.

Another example is ISO/IEC TR 29138-1:2009, a technical report that consolidates prevalent accessibility requirements for users with disabilities. This document offers an enumeration of comprehensive requirements across approximately 14 distinct areas, functioning as a reference for standards developers. The report underwent revision in 2018 to accommodate users' needs in the current technological landscape.

In the realm of web accessibility, JTC 1's most noteworthy achievement was the formalisation of W3C guidelines as international standards. In 2012, ISO/IEC JTC 1 published ISO/IEC 40500:2012, which serves as an exact adoption of the W3C Web Content Accessibility Guidelines (WCAG) 2.0. WCAG 2.0 comprises twelve guidelines, along with detailed success criteria organized under four principles: Perceivable, Operable, Understandable and Robust. By achieving the ISO/IEC 40500 status, these guidelines can be referenced more effectively as national standards worldwide. (Note: While W3C released WCAG 2.1 in 2018, the ISO version remains at 2.0 as of 2025, with expectations that WCAG 2.2 and subsequent versions will also be adopted as international standards.)

The interoperability of hardware and assistive technology interoperability has been a key concern for JTC 1. The ISO/IEC 29136:2012 standard, entitled "Accessibility of Personal Computer Hardware", establishes methodologies for the design of hardware interfaces – such as keyboards, mice and displays – to ensure accessibility for persons with disabilities (for instance, using Braille displays, tactile markers, force thresholds). The ISO/IEC 13066 series pertains to the compatibility between operating systems and assistive technologies (including screen readers and specialized mice). Part 1 (2011) delineates the requirements for interoperability, while the accompanying annexes offer environment-specific guidelines for Windows, Linux and other systems.

Furthermore, the ISO/IEC 24751 series (2008–2013) has established standards for articulating individual accessibility preferences within personalized learning environments, thereby contributing to the definition of the concept of "AccessForAll" in the field of education.

Recently, standards have been developed in response to the emergence of artificial intelligence (AI) and the diversification of digital content. ISO/IEC 30071-1:2019 provides a comprehensive code of practice for incorporating accessibility considerations throughout the ICT product and service development process. It systematizes the approach by which development organisations can define accessibility objectives and effectively implement and verify them.

Other technical committees under the International Organisation for Standardisation (ISO) have also published related standards. For instance, ISO 21801-1:2020, entitled "Cognitive Accessibility Guidelines", delineates principles for the design of products and services intended for users with cognitive disabilities, making it a significant reference in AI-based service design. Additionally, the International Electrotechnical Commission (IEC) has released IEC 63008:2020, which specifies methods for testing accessibility in household appliances. This aims to improve the usability of smart home and Internet of Things (IoT) devices for persons with disabilities.

Efforts have also been undertaken to incorporate accessibility considerations into the standardization process itself. ISO/IEC and ITU have collaboratively published ISO/IEC Guide 71:2014, titled "Guide for addressing accessibility in standards". This guide provides practical recommendations to standard developers regarding the systematic incorporation of the needs of people with disabilities, older populations, children and others into all standards. It advocates for strategies and objectives concerning accessibility, including the principle to "consider universal access from the outset of design". Consequently, due to the dissemination of this guide, a greater number of new ISO and ITU standards now encompass accessibility clauses, thereby reinforcing the global trend towards digital inclusion through standardization.

7.3 W3C

In 1997, W3C launched the Web Accessibility Initiative (WAI), continuing to develop standards aimed at enhancing accessibility and inclusiveness for people with disabilities. Early on, it established foundational guidelines such as WCAG 1.0 (1999), ATAG 1.0 (2000) and UAAG 1.0 (2002), which set principles for web content, authoring tools and user agents (browsers). In 2008, W3C adopted WCAG 2.0 as a formal recommendation, restructuring the guidelines, and later, in 2015, updated them with ATAG 2.0 and UAAG 2.0. WCAG 2.1 (2018) addressed mobile environments and the needs of users with cognitive, learning and low-vision disabilities, while WCAG 2.2, released in 2023, further expanded the success criteria.

In parallel, the WAI-ARIA (Accessible Rich Internet Applications) specification evolved to support dynamic web content – initially released as version 1.0 in 2014, updated to 1.1 in 2017 and again to 1.2 in 2023. W3C is also developing WCAG 3.0, a next-generation set of guidelines intended to encompass not only web content but also applications, authoring tools, publishing and other digital technologies within a more inclusive framework.

WCAG 2.2, published in October 2023, added nine new success criteria to the existing WCAG 2.1 framework, strengthening accessibility for users with cognitive impairments, low vision and those who rely on keyboard navigation. Notable additions include clearer focus indicators, alternatives to drag-and-drop actions, minimum touch target sizes and simplified authentication methods. These enhancements maintain backward compatibility while improving support for mobile and modern interfaces.

Meanwhile, WCAG 3.0 remains in draft form and proposes a fundamental shift from binary pass/fail criteria toward a more flexible, tiered conformance model (Bronze, Silver, Gold). It expands the scope to cover a broader range of digital content, including applications, PDFs and IoT devices. WCAG 3.0 emphasizes user-centred accessibility, technical neutrality and measurable outcomes, aiming to improve inclusiveness and real-world effectiveness.

8 Conclusions

As explored above, various international organizations have been actively developing standards related to accessibility and inclusiveness for people with disabilities. While each organization brings its unique perspective – rooted in its founding mandate and specialized focus – the convergence of emerging technologies is fostering new opportunities for synergy. As areas of work increasingly intersect, there is greater potential for collaboration, alignment, and mutual reinforcement. This evolving landscape has already produced impactful results, such as the adoption of W3C Web Content Accessibility Guidelines (WCAG) as an ISO/IEC standard, and the joint development of the "Guide for addressing accessibility in standards" by ISO/IEC and ITU. These successes demonstrate the value of working together to advance common goals and ensure inclusive outcomes.

Initially centred around web accessibility, the scope of accessibility-related standardisation has since expanded to encompass a wide range of technologies and services, including mobile devices, software, digital services, kiosks and assistive ICT. More recently, new standards are being developed to ensure that no group is excluded from the benefits of emerging technologies such as smart cities and artificial intelligence, particularly in promoting better mobility and equitable access to information.

This section provides a summary of the report and outlines potential next steps, particularly in relation to standardization, research, and cross-sectoral collaboration, both within ITU and beyond.

The report, after an introduction, defines the public and private stakeholders in Section 2. Along with industry, communities such as people with disabilities NGOs and academic and standardization bodies are essential collaborators for building inclusive virtual spaces. Section 3 defines the many interactions required to access virtual worlds. The fourth section focuses on tools and technologies towards the interactions and Section 5 defines the many accessibility services. Section 6 presents some use cases demonstrating all the key elements of the human-centric design and the practical benefits of digital inclusion in virtual settings. These encompass various aspects of human activity, including work, leisure, sport and art. The last section provides an essential overview of standardization efforts by the three key international bodies – ITU, ISO and W3C – highlighting their contributions and the growing convergence towards comprehensive digital inclusion through a unified approach.

Looking ahead, selected elements of this technical report may serve as the basis for proposed New Work Items under ITU-T Study Group 20, particularly within Question 8 on Human-centric digital services enabled by the Internet of Things and smart sustainable cities and communities, thereby contributing to the development of robust and effective international standards.

About the Global Initiative on Virtual Worlds and AI – *Discovering the Citiverse*

Launched by ITU, UNICC, and Digital Dubai, the Global Initiative on Virtual Worlds and AI – *Discovering the Citiverse* is a multistakeholder platform dedicated to shaping the next generation of AI-powered virtual worlds.[52]

These immersive digital environments are transforming how people live, learn, govern, and interact. The Initiative ensures that AI-powered virtual worlds evolve in ways that are inclusive, interoperable, and human-centric–and that they help deliver on the Pact for the Future and its Global Digital Compact.

Serving as a neutral and action-oriented platform, the Initiative brings together cities, governments, UN agencies, private sector companies, academia, and civil society to collaboratively shape the responsible development and deployment of these technologies.

The Initiative advances its mission through three strategic pillars, each supported by dedicated tracks that address the most urgent challenges and promising opportunities in AI-powered virtual worlds. This comprehensive structure enables the Initiative to deliver both high-level global guidance and practical implementation in cities worldwide.

The Initiative is supported by over 70 international partners.

For more information, please visit: https://www.itu.int/metaverse/virtual-worlds/.

Meet the Champions

Champions are entities that demonstrate leadership by providing financial contributions in support of the Initiative. This may include funding for events, challenges, research outputs, communication activities, trainings, travel grants, or other related efforts.

Ministry of Internal Affairs and Communications

دبي الرقمية DIGITAL DUBAI

Meet the Founding Partners

Founding Partners are the organizations that launched the Initiative. They serve as the core convening entities and contribute to shaping its long-term vision. The Founding Partners are:

ITU

UNICC

دبي الرقمية DIGITAL DUBAI

Meet the Supporters

Supporters are organizations that have expressed endorsement of the Initiative and actively participate in its activities. This includes, but is not limited to, participation in tracks, contribution of use cases, co-organization of events, provision of expertise, or public advocacy of the Initiative.

REPUBLIC OF KENYA MINISTRY OF INFORMATION, COMMUNICATIONS AND THE DIGITAL ECONOMY

ETRI

Stanford University

IDC

References

1 https://www.w3.org/WAI/standards-guidelines/wcag/

2 https://www.iso.org/standard/27001

3 https://oauth.net/2/

4 https://w3c.github.io/webauthn/

5 https://www.nist.gov/privacy-framework/nist-sp-800-162

6 https://www.iso.org/standard/27001

7 https://www.iso.org/standard/86013.html

8 https://www.iso.org/standard/70721.html?utm_source=chatgpt.com

9 https://www.iso.org/standard/30824.html

10 https://www.iso.org/standard/73227.html

11 https://standards.ieee.org/ieee/3079/7283/

12 https://www.w3.org/immersive-web/

13 https://www.iso.org/standard/73227.html

14 https://standards.ieee.org/ieee/3079/7283/

15 https://www.iso.org/standard/84636.html

16 https://www.iso.org/standard/41521.html

17 https://www.iso.org/standard/61362.html

18 https://www.iso.org/standard/54106.html

19 https://www.iso.org/standard/25508.html

20 https://www.iso.org/standard/82562.html

21 https://www.iso.org/standard/41521.html

22 https://www.w3.org/TR/webrtc/

23 https://www.rfc-editor.org/info/rfc3550

24 https://www.iso.org/standard/57359.html

25 https://www.iso.org/standard/41521.html

26 https://gdpr-info.eu

27 https://www.iso.org/standard/70331.html

28 https://www.iso.org/standard/80392.html

29 https://www.iso.org/standard/41521.html

30 https://www.iso.org/standard/41521.html

31 https://www.iso.org/standard/64840.html

32 https://www.iso.org/standard/64840.html

33 https://www.iso.org/standard/63500.html

34 https://www.iso.org/standard/41521.html

35 https://www.iso.org/standard/77520.html

36 https://www.iso.org/standard/64840.html

37 https://www.iso.org/standard/41521.html

38 https://www.iso.org/standard/81896.html

39 https://www.iso.org/standard/78176.html

40 https://www.iso.org/standard/41521.html

41 Ryan Golden, "BofA Banks on Virtual Reality to Train Workers for the 'Moments of Truth", HR Dive, 16 March 2021, https://www.hrdive.com/news/bofa-banks-on-virtual-reality-to-train-workers-for-the-moments-of-truth/596780/. Katherine Doherty, "Bank of America Is Using AI and Metaverse to Train New Hires", Bloomberg, 13 July 2023, https://www.bloomberg.com/news/articles/2023-07-13/bank-of-america-is-using-the-metaverse-ai-to-train-its-hires. Bank of America Press Release on VR Training Launch (October 2021), available at: https://newsroom.bankofamerica.com/content/newsroom/press-releases/2021/10/bank-of-america-is-first-in-industry-to-launch-virtual-reality-t.html. PR Newswire Announcement of VR Training Rollout, available at: https://www.prnewswire.com/news-releases/bank-of-america-is-first-in-industry-to-launch-virtual-reality-training-program-in-nearly-4-300-financial-centers-301395210.html.

42 The title "Crip Sensorama" is a humble nod to the physical device "The Sensorama" (often framed as the earliest VR technology) built in 1962 by filmmaker Morton Heilig. The use of the word "Crip" is a reclaiming of the word "Cripple" as a marker of proud and defiant identification of disabled bodies, often celebrated in disability circles.

43 Diogo Furtado, Renato Alexandre Ribeiro, Manuel Piçarra, Letícia Seixas Pereira, Carlos Duarte, André Rodrigues, and João Guerreiro. 2025. Designing and Evaluating a VR Boxing Experience with Blind People. In Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25). Association for Computing Machinery, New York, NY, USA, Article 1246, 1–17. https://doi.org/10.1145/3706598.3713374.

44 https://www.itu.int/ITU-T/recommendations/rec.aspx?rec=9017

45 https://www.itu.int/ITU-T/recommendations/rec.aspx?rec=13661

46 https://www.itu.int/ITU-T/recommendations/rec.aspx?rec=14341

47 https://www.itu.int/ITU-T/recommendations/rec.aspx?rec=13662

48 https://www.itu.int/ITU-T/recommendations/rec.aspx?rec=13571

49 https://www.itu.int/ITU-T/recommendations/rec.aspx?rec=13858

50 https://www.itu.int/en/ITU-T/studygroups/2025-2028/20/Pages/default.aspx

51 https://www.itu.int/en/ITU-T/studygroups/2025-2028/20/Pages/default.aspx

52 https://www.itu.int/metaverse/virtual-worlds/

For more information,
please contact: virtualworlds@itu.int
Website: https://www.itu.int/metaverse/virtual-worlds/

ISBN 978-92-61-41471-9

Published in Switzerland Geneva, 2025

Photo credit: Adobe Stock