The Key to an Open, Functional, and Interoperable Metaverse

Anita Havele
Web3D Consortium, USA

Don Brutzman
Naval Postgraduate School, USA

William Benman
IVN Networks

Nicholas F. Polys
Virginia Tech

https://doi.org/10.15221/22.27

Abstract: The term 'Metaverse' has taken on new interest recently, appearing prominently in the marketing materials of a number of large technology companies. Indeed, many have attempted, or are attempting, to co-opt it for their own purposes, which has resulted in a great deal of confusion among producers and consumers in the marketplace. With this paper, the Web3D Consortium seeks to address this confusion by exploring the history of the 'Metaverse', provide a workable definition of the term 'Metaverse', and provide a vision for its sustainable, cooperative construction into the future. We believe that all the technologies are in place to fulfill the vision of an open, equitable, and ubiquitous information space. What remains are the key issues that have kept the Metaverse from manifesting the last two decades: poor user experience and poor corporate cooperation.

Keywords: Metaverse, Open Standards, Interoperability, Web3D, 3D Graphics

Corresponding author: Anita Havele. Email: anita.havele@web3d.org.

Introduction and Background

While the media struggles with an understanding and definition of the term and future of the "Metaverse', it is fitting to reflect on the history of the Metaverse that has brought us to this moment in time. One can easily argue that 'Metaverses' have existed since the 90s and are now in their fourth generation of development Howard, 2022, Ravenscraft, 2022, Cozzi and Marc, 2022, Ghent, 2021 Pennington, 2022, Parisi, 2021 Dionisio et al., 2013. Since the early days of the World Wide Web (WWW) and HTML, the Web 3D community has built technologies that link 3D worlds and assets interactively across the network. In the 1990s we were on 28.8 modems; there were no affordable GPUs or headsets, and virtual environments were an aspiration. Still, the vision of shared digital worlds and VR was established as a compelling motivator for the community.

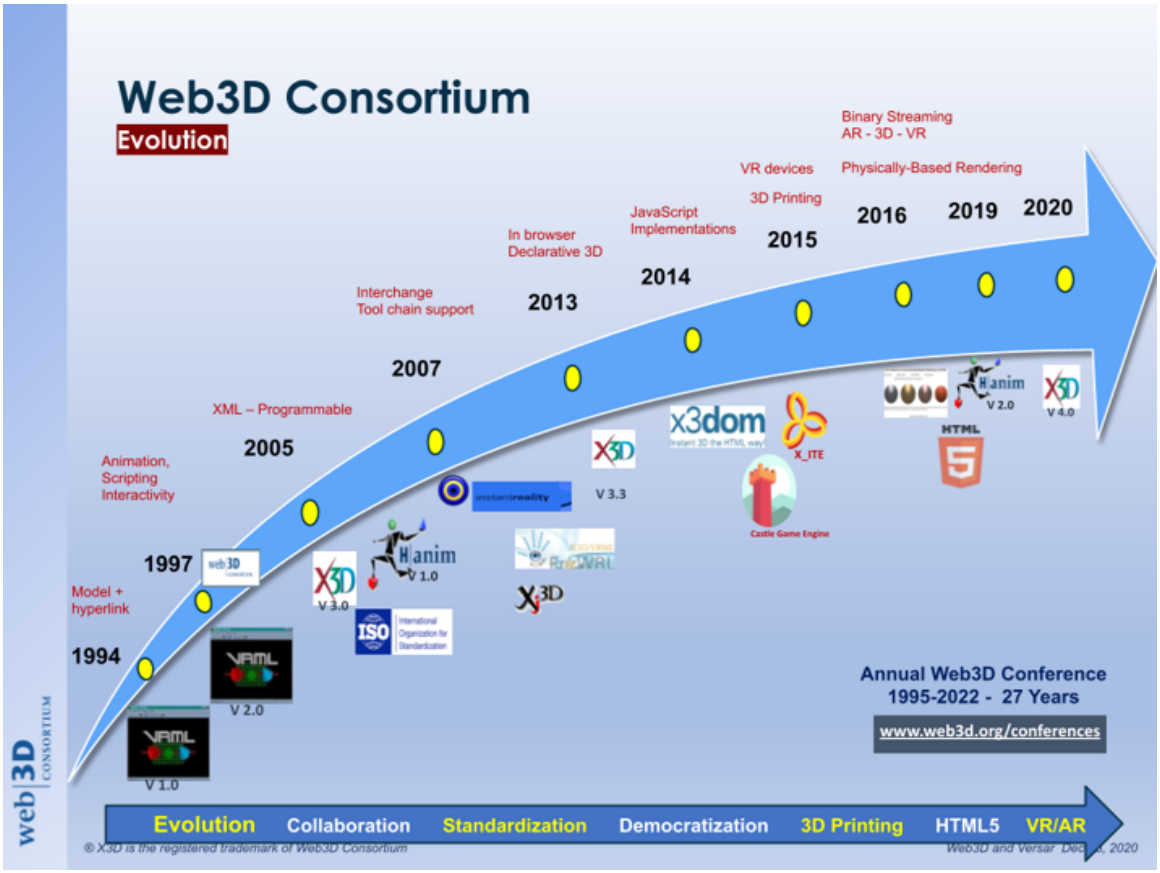

Figure 1. ***The longitudional development of the ISO-IEC standards VRML and X3D.***

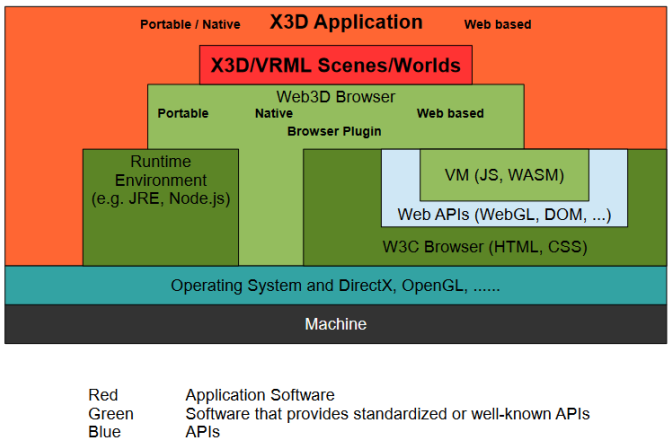

Figure 2. ***Example X3D APIs and Applications stack from a software perspective***

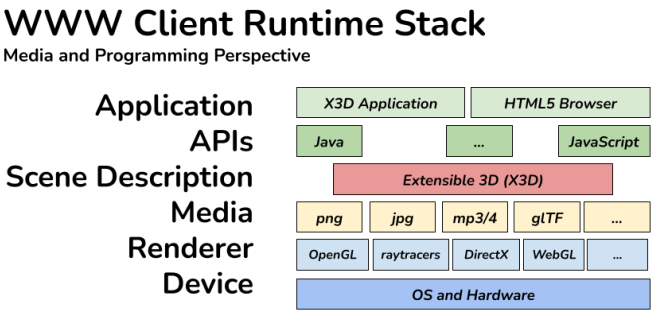

Figure 3. ***Example X3D Applications stack from a media-type perspective***

To support the description and sharing of these interactive 3D worlds, the community established the not-for-profit Web3D Consortium™ (Web3D.org) and its members have been aligning Web standards around 3D graphics, animation, and interaction for decades ever since. Many innovative and powerful capabilities that we take for granted today have emerged on top of this stable foundation. Figure 1 shows the evolution of open Web3D™ Standards since 1997 VRML. ISO-IEC Standards provide a strong formal process. The Web3D Consortium created the ISO-IEC Standards of X3D (second generation VRML) and HANIM (for humanoid avatars). Through collaborations and liaison relationships with W3C and other Standards Development Organizations (SDO), Web3D Consortium is developing the suite of 3D technologies for the open Web stack (Figure 2).

Standards and interoperability support the long-term strategy toward the durability of 3D content. Whether it is a nuclear power plant, scientific products, or an institutional facility, the data durability requirements far exceed Silicon Valley's lifecycles. For example, the portability and durability of the ISO-IEC interactive scene graph has enabled an unmodified 1998 Web3D (VRML) world to run faster-than-ever in 2022. This content interactive 3D content still runs on modern systems, even with multiple GPUs, stereo rendering (27.6 million pixels), and immersive user tracking in the Virginia Tech Visionarium Lab's HyperCube (CAVE).

Meanwhile, others focused on capture and monetization of user traffic in online 3D worlds. From There, to Olive, VastPark, Avaya, companies set up proprietary virtual worlds and real-estate for people to meet and collaborate in. These worlds and the companies that built them are gone N. F. Polys, 2011, N. F. Polys, 2017. The Second Life platform has become another microverse. Such walled gardens indicate their assumption of the Highlander Theory: i.e that "There can be only one". However, the Web is an ecosystem and one must play well with others to survive. Google's Lively and more recently, Poly, are examples that bankruptcy is not the only reason for companies to discontinue various Metaverse product lines.

Current Status

Many are describing specific games and virtual worlds as 'The Metaverse'. As we have discussed, many years of innovation have shown us the possibilities of virtual worlds. Games, social networks, and augmented reality applications have driven exciting growth and expansion. However, these virtual worlds are still generally restricted for use within their own proprietary ecosystem, i.e. 'walled gardens', where servers control and admit users, content, and services and are not interoperable with other virtual worlds. As such, movement between worlds is not seamless. Instead, movement between worlds requires signing out of one world, thereby dropping out of 3D space into a conventional 2D Web page, then signing into another 3D world. Thus, from a user perspective, there is little to no 'interoperability' or 'accessibility'.

This begs the questions:

1. User Experience:

    - How will we achieve accessibility? ... allowing us to move ourselves and our content seamlessly between real, virtual, and augmented worlds?

    - How will we ensure security and provide for user-control of personally identifiable information (PII) or even our own health data?

2. Interoperability:

    Is the Metaverse just a single walled garden, shared game or social environment? Or will the Metaverse be an open, unified space suitable for critical business, education, ecommerce, entertainment, e-government, social applications and industry applications? How?

The Web3D Consortium and its members are committed to addressing these limitations in a comprehensive and sustainable way through our proven standardization process and long-running liaison partnerships with other standards development organizations (SDOs) including W3C, Khronos, OGC, HL7, and DICOM.

Definition and Future of the Metaverse

We believe that a fully successful 'Metaverse' will not be a collection of separate "walled gardens" or 'microverses'. Rather, we see it as a constellation of connected multi-dimensional realistic and/or fantasy computer generated (i.e. virtual) or augmented worlds in which people will work, learn, play, buy, sell, communicate, collaborate, interact and travel, pursuing their tasks without the constraints of

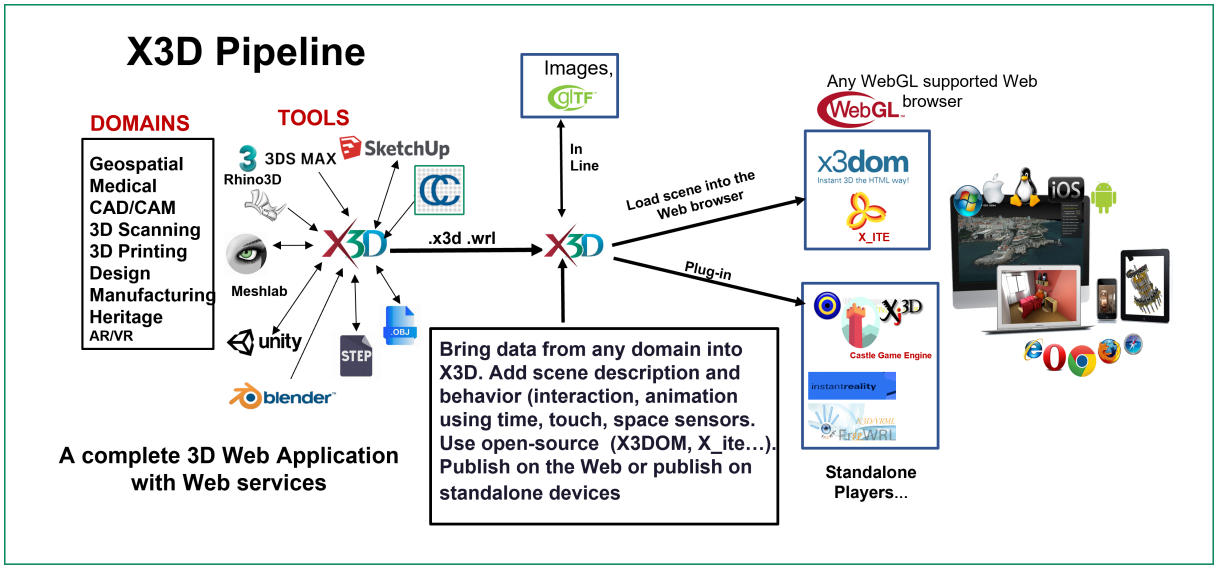

Figure 4. *X3D functional support for the publishing ecosystem*

physical space and time. We believe that a 'Unified Metaverse' will require all of the above as well as the ability to move between 'microverses' without dropping out of 3D mode. Such unification will require interoperability and this interoperability will require an open Web standard and a technology to facilitate seamless movement between worlds.

Therefore, in our view, the 'Metaverse' will emerge as a property of the current primarily two-dimensional Worldwide Web (WWW) with interconnection and interoperability between online, networked microverses created, as is the Web, by individuals, corporations and institutions. Interconnection between microverses will be provided by the World-Wide Web itself.

The Extensible 3D Standard (X3D)

Interoperability between interconnected microverses can only be provided through Standards support across microverses, which allows users to move seamlessly from one microverse to another - ideally while maintaining their selected and secured identity. The Web3D Consortium has developed a standard that is ideal for this purpose. The standard is Extensible 3D (X3D). The X3D Standard is an open technology that comprises many of these imperatives and priorities, while leveraging on other industry standards including WWW, WebGL, and glTF. X3D® is the royalty-free, open ISO-IEC Standard for the publishing, viewing, printing and interactive communication of 3D virtual worlds on the Web (Figure 4). These virtual multimedia and data-driven worlds can be experienced synchronously or asynchronously at a WWW address. This has been a reality of virtual reality and Web3D for over 20 years. To learn more, many X3D Example scenes and archives are available online "X3D Online Examples and Archive", 2022. The X3D Version 4 Overview "X3D design criteria and specification publication", 2022 includes design criteria and specification publication links.

Extensible 3D (X3D) has a rich vocabulary for describing 3D worlds across many disciplines and vertical industry applications, for example, CAD, GIS, AR/VR, 3D Printing and Scanning, shaders, volume

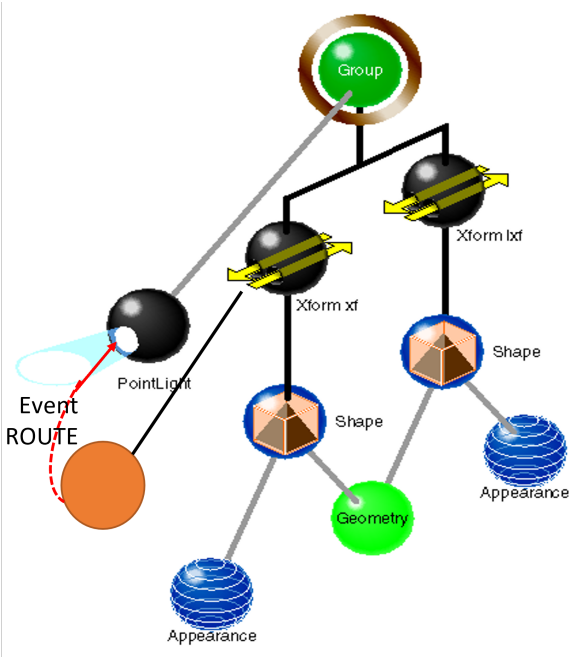

Figure 5. ***An X3D scene graph illustrated to show both the transformation graph and the behavior graph: touching a shape turns on the light***

rendering. As a stable and capable digital representation, X3D enables serious 3D applications across multiple industries. As new technologies and features stabilize, Standards evolve. The latest X3D® version 4 (X3D4) is a major upgrade that provides close support for HTML5, Khronos glTF Physically-Based Rendering (PBR), and the WebAudio API Lakka et al., 2020, Lakka et al., 2021. X3D4 includes multiple file encodings and programming language bindings to support a variety of implementations. An extensible, greatest-common denominator representation for 3D interactive graphics information, X3D has adapted, specialized, and grown over the years.

In addition, The X3D scene graph (Figure 5) can be expressed in Semantic Web Ontologies Flotyński et al., 2019, Flotyński et al., 2020, which makes it powerful for transcoding 3D assets and interactions between Microverses' walled proprietary gardens and formats. Finally, the scene graph can be manipulated by programs written in different languages: JavaScript, Java, Python, C#, and C++; these APIs are specified in X3D's suite of ISO-IEC ratified standards (Figure 6).

The Web3D Consortium is committed to supporting the creation of a stable, equitable, and evolving Metaverse. X3D is, by its nature, Extensible; meaning that the language allows for custom nodes and content that may be outside the ISO-IEC Standard. For example, a Microverse may use a special geometry or appearance, or metadata. By describing such nodes and attributes in Profiles (Figure 7) within the X3D scene graph, novel and innovative content can still seamlessly move in real-time across 3D applications, networks, and Web services. We believe this interoperability is an essential ingredient for success. Consider, if an asset such as a 3D model of a car is used in the Metaverse, should its behavior be consistent? For example, the tires and steering wheel work on a specific axis and are controlled by the driver's interaction; touching the door handle opens it predictably, touching a dashboard button turns on the headlights. X3D thus enables rich, interactive content to be re-used across worlds.

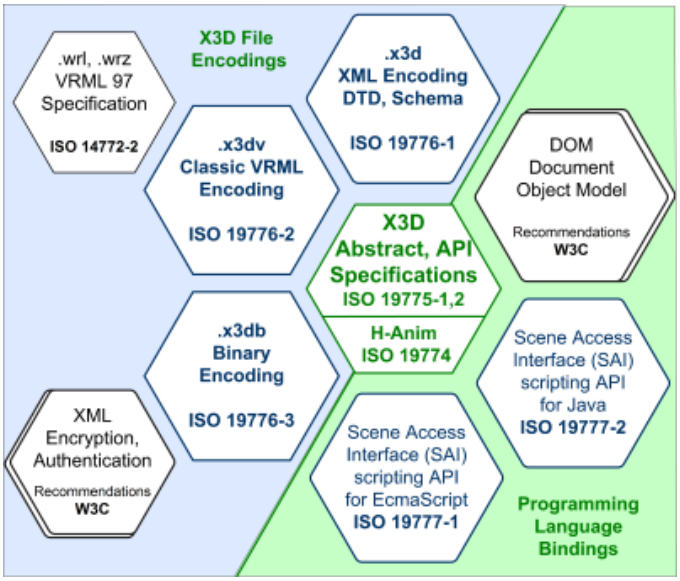

Figure 6. ***The ISO-IEC family of X3D standards.***

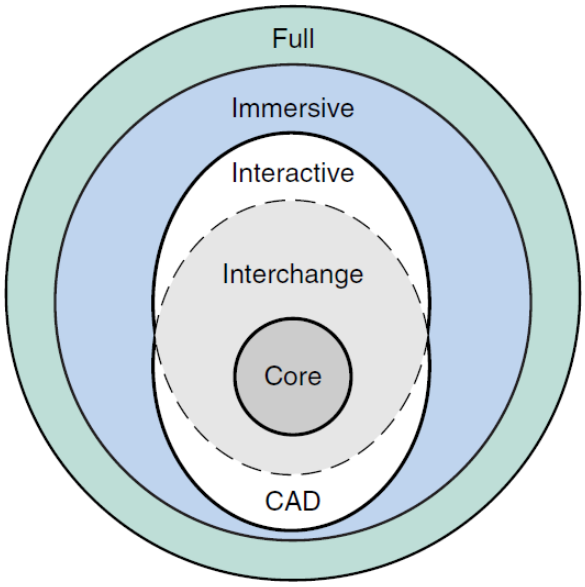

Figure 7. ***Profiles of X3D functionality permit componentized complexity in model sophistication and browser capabilities.***

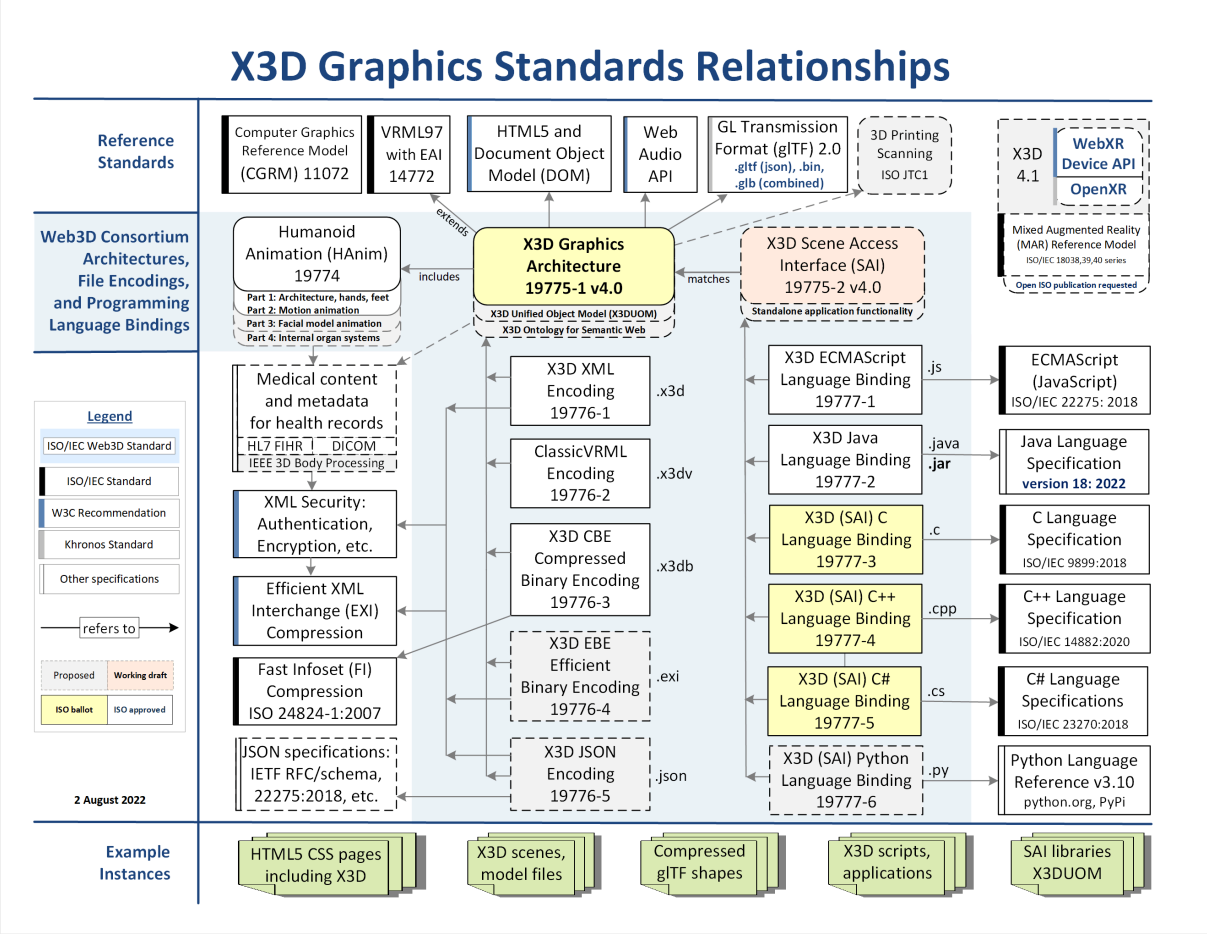

Figure 8. ***X3D Standards ecosystem includes multiple file encodings and programming-language bindings***

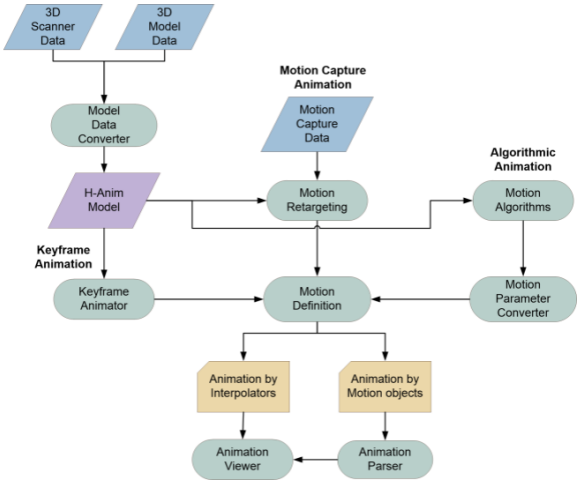

Figure 9. ***Scope of Humanoid Animation Part 2.***

Understanding the role and goal of each technology, its community, specifications, and Standards are essential for a successful approach and strategy for interoperability. For example, the design goal of glTF 2.0 is to be the JPEG of 3D. In the same way, X3D and VRML are the HTML of 3D. glTF are assets that are composed into an X3D scene or application, at the Web and Application layer N. Polys, 2019. A Features Comparison between X3D4 and glTF2 shows interoperability between glTF and X3D Group, 2022. X3D's strength lies in composing interactive scenes (via its Scene Graph) and connecting them with higher-level logic, APIs, and services. Figure 8 illustrates X3D's position in the Standards ecosystem.

Building durable, non-proprietary worlds is just one piece of the puzzle. These standards are openly published and, being extensible, have provided the basis for decades of innovation, not just in this conference series. A quick search on Google Scholar returns for 21,600 documents for X3D and for 72,300 for VRML; at Semantic Scholar, they have indexed 2,140 papers with X3D in the title and 10,100 with VRML in the title. The textbook X3D: Extensible 3D Graphics for Web Authors Brutzman and Daly, 2010 explains basic X3D model design up through the X3D Immersive Profile. Corresponding course videos and slidesets with a full set of example X3D scenes demonstrate best-practice modeling techniques for Web authors.

Avatar Strategy: Open, Interoperable, Personal, Secure

For an avatar strategy that is open, Interoperable, Personal and Secure we will need:

- Multiple levels of detail (LOD) to progress avatar fidelity for highly scalable shared worlds
- Metadata for identity, secure digital signature and even encryption. Security applied to the entire model or selective portions, as appropriate.
- Efficient, private, and secure 3D scans and motion-capture (mocap) data for full avatar realism .
- Avatars with sufficient detail to allow 'standard' animations based on realistic rotations and translations to be shared between 'standard' human models.

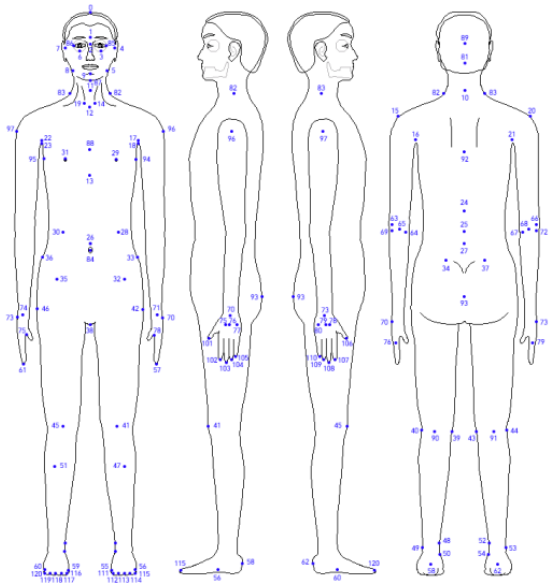

Figure 10. ***Humanoid Animation Feature Points.***

WebD Consortium's standard, X3D Humanoid animation (HANIM) is specified in ISO/IEC 19774 which provides complete normative and informative detail to specify an abstract human form. The specification defines the skeleton arrangement of joints, segments, and sites exhibiting four levels of articulation along with the normative initial or default skeleton pose. HAnim 2.0 specification includes complete expression for joints and segments making up full medical anatomy of humans. HANIM intends to demonstrate whatever avatar definitions are agreed upon by the industry. We expect to show full matching. Implementation + evaluation to show what avatar designs work best. Time has come for serious implementation and further standardization!

X3D Body Tech and Medical

Developing an open interoperable standard for the representation of human anatomy based on input from a wide variety of imaging modalities is very important. If interoperable, researchers and medical professionals can take data from many different types of equipment and fuse them into a coherent 3D data set that can be used both for patients and medical professionals. Medical and humanoid applications can benefit from real time 3D visualization.

Many X3D HANIM use cases are now possible using data surfaces, materials, and Hierarchies information from body scans for clothing, Dermatology, Plastic surgery and more. Imaging and volume rendering is another alternate form of 3D data representation compared to the traditional polygonal form to implement some of the higher-complexity representations in a model. With Ontology for Semantic Web and Metadata FMA and SNOMED integration is also available.

Requirements

If the Metaverse is to make an impact on our digital lifestyles, there must be compelling reasons to be there:

- User activities should not be fragmented across multiple client applications and devices
- User experience should not be interrupted when traveling between microverses
- Users have agency and responsibility for world creation, interaction, deletion
- Users feel safe in the protection of their personal space and identity

User Experience

Safety and security of immersive technologies are a critical part of how people might engage in Metaverse interactions over the Web. How are people protected? How are accessibility concerns addressed when everyone has some potential for access-impaired interaction? In collaborative systems, one should be able to reliably opt-out if they are concerned about privacy. For example, there could be global settings on the browser or OS for privacy that take advantage of the World Wide Web security stack (as X3D does).

The functional design of the X3D Architecture has multiple safeguards against malicious model content, and offers few threat vectors to bad actors. Many lessons learned from HTML are included. Special constraints for 3D scenes are testable and implementable by X3D browsers, for example the restriction that X3D models cannot recursively load themselves, either directly or indirectly. These preventative practices apply to all of the various X3D file formats (XML ClassicVRML JSON JavaScript Java Python C/C++/C#, compressed, etc.) providing a secure basis throughout X3D worlds.

The .x3d file format encoding of the X3D scene graph is XML based, which means that the full power of Web Security can be applied to X3D Graphics models. Open security by design, as specified by W3C's XML Security Recommendations, provides the strongest approach for secure Web documents used internationally on the Web. Capabilities include Digital Signature for authentication and Encryption for privacy, applied independently or together on any element in the scene.

An Example

Considering our requirements above, we can concretely consider how X3D supports them. With X3D, 3D worlds and assets live at url addresses; they may be files or they may be created dynamically as a service and delivered using X3D MIME types. Leveraging W3C Standards, X3D has support for marshalling assets across the network and hyperlinking between worlds (through urls). By describing Web Metaverses with interactivity, animation, and lighting in X3D, authors and users can have access to the scene graph through any browser supporting HTML5 and Javascript.

In the case of security, we consider both network and data -centric approaches. Certainly in the WWW, network authentication is widely used to give tiered access to various controlled backend services. Through the data-centric approach, we focus on encrypting the content itself. X3D's XML encoding supports use cases where some pieces of the 3D scene are encrypted and only visible to authenticated users (W3C's XML Security). For example, using the Level-Of-Detail (LOD) or Switch mechanism in the X3D scene graph enables multiple resolutions of the same object to co-exist. While I might use a high-resolution body scan as my avatar in the Metaverse, only my doctor or spouse should be able to decode that file and see it; for work, perhaps, I have a lower resolution that my co-workers can authenticate to see; for social sites, perhaps it is a 'cartoon' version that is published for participation in public spaces. One

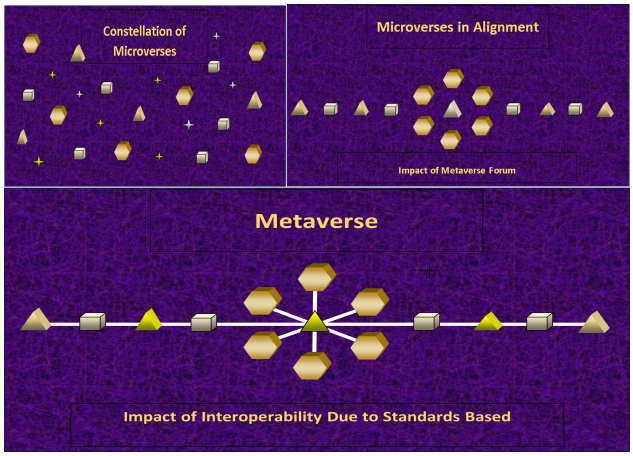

Figure 11. ***Standards-based interoperability evolution: (a) Constellations of Microverses, (b) Microverses in Alignment, and (c) eventual interoperable Metaverse capabilities.***

can extend this thought experiment pattern to secure the functionality described in any X3D Component and Profile.

Future

We have shown how X3D, the international ISO-IEC Standard, provides the essential ingredients to scale Metaverse functionality to WWW scale through its integration of both declarative assets (Linked Data) and language runtimes. The ISO-IEC standardized X3D scene graph can be the greatest-common denominator to catalyze the WWW Metaverse: the "Presentation Layer" that supports interactive Web and Web3D applications across a host of real and compelling use cases.

X3D standards are platform-agnostic and support many 3D models formats and market domains. In the last two and half decades, established and new graphics features have all been map-able to X3D and its extensibility mechanism, often to their eventual standardization. The Web is X3D's platform. Web3D Consortium members develop and extend a proven technology that provides for the interoperability between interactive 3D worlds and to enable an open, unified Metaverse. On the solid foundations of Web3D, W3C, and SDO standards we can enable users to build and share their part of the Metaverse and move seamlessly in and out of specialty microverses.

Our vision for the Metaverse is one that is open and unified with seamless movement between diverse worlds without leaving the immersive 3D environment experience (Figure 9). We have argued that the true value of a Metaverse lies in its seamless user experience across content and systems. Figure 9a (top-left) shows the constellation of microverses that is representative of where we are today with multiple microverses on diverse platforms. Each microverse is a microcosm in and of itself, with no option to move between microverses. Figure 9b (top-right) shows the anticipated alignment for the creation of the Metaverse, expected to flow from guidance from standards development organizations, co-operations from the corporate along with input from the user community. Success in creation of a standard(s) for the Metaverse and adoption of the standard by the constellation of microverses will lead to the creation of the Metaverse depicted in Figure 9c (bottom-center). Figure 9c Shows a Metaverse in with full interoperability between microverses pursuant to a common standard. This interoperability is expected to enable users

to move between microverses without exiting 3D mode while retaining their identity and data. This is depicted as interconnections between microverses by which it should be understood that any user in any microverse can move to another, if the user is a member or citizen of both microverses.

When the user experience is prioritized, it will drive the relevant standard(s) to adoption. We believe that progress toward this vision of the Web-enabled Metaverse will happen sooner with corporate cooperation. When we say cooperation, we mean CO-OPERATION through the formal processes and transparent activities of standardization. We are not recommending one large corporation adopting another large corporation's proprietary format, which is more about movies than interoperable interactive 3D experiences. While the current Metaverse has been building for many years through WWW standards, we see some positive developments in the awareness and adoption with the newly formed Metaverse Standards Forum that has over 1,200 companies participating. These discussions are new and are rapidly evolving. We believe that through education and the sharing of knowledge, we can achieve our common goals more quickly. As we describe in this paper, X3D provides a unique value proposition to quickly integrate many existing systems in a proven framework of internationally-ratified Standards.

References

Brutzman, D., & Daly, L. (2010). *X3d: Extensible 3d graphics for web authors*. Elsevier.

Cozzi, P., & Marc, P. (2022). Open metaverse podcast. https://cesium.com/open-metaverse-podcast/

Dionisio, J. D. N., III, W. G. B., & Gilbert, R. (2013). 3d virtual worlds and the metaverse: Current status and future possibilities. *ACM Comput. Surv.*, *45*(3). https://doi.org/10.1145/2480741.2480751

Flotyński, J., Brutzman, D., Hamza-Lup, F. G., Malamos, A., Polys, N., Sikos, L. F., & Walczak, K. (2019). The semantic web3d: Towards comprehensive representation of 3d content on the semantic web. *2019 International Conference on 3D Immersion (IC3D)*, 1–9.

Flotyński, J., Malamos, A. G., Brutzman, D., Hamza-Lup, F. G., Polys, N. F., Sikos, L. F., & Walczak, K. (2020). Recent advances in web3d semantic modeling. *Recent advances in 3D imaging, modeling, and reconstruction*, 23–49.

Ghent, E. (2021). Why the metaverse needs to be open making virtual worlds as interconnected as the internet will be tough. *Spectrum*. https://cesium.com/open-metaverse-podcast/

Group, X. W. (2022). Graphics scope comparison: Gltf and x3d. https://www.web3d.org/specifications/FeaturesComparisonX3D4glTF2.pdf

Howard, B. (2022). Remembering vrml: The metaverse of 1995. *How to Geek*. https://www.howtogeek.com/778554/remembering-vrml-the-metaverse-of-1995

Lakka, E., Brutzman, D., Puk, R., & Malamos, A. G. (2020). Extending x3d realism with audio graphs, acoustic properties and 3d spatial sound. *The 25th International Conference on 3D Web Technology*, 1–9.

Lakka, E., Papadaki, M., Brutzman, D., Puk, R., & Malamos, A. G. (2021). X3d audio graph for the consistent declarative representation of the w3c audio api. *The 26th International Conference on 3D Web Technology*, 1–5.

Parisi, T. (2021). The seven rules of the metaverse. https://medium.com/meta-verses/the-seven-rules-of-the-metaverse-7d4e06fa864c

Pennington, A. (2022). An open metaverse is about more than just interoperability — it's about accessibility. *Amplify*. https://amplify.nabshow.com/articles/ic-an-open-metaverse-is-more-than-just-interoperability-its-about-accessibility/

Polys, N. (2019). Integrating x3d and gltf: How to play together well on the www. https://www.web3d.org/blog-integrating-x3d-and-gltf

Polys, N. F. (2011). Recall the lost frontiers of virtual worlds. *Communications of the ACM, 54*(5).

Polys, N. F. (2017). Of standards and herrings: Tales of technology and tumult. *COMSOC MMTC Communications Frontiers*. https://site.ieee.org/comsoc-mmctc/files/2016/04/MMTC_E-Letter_Jan_2017_singlecolumn_v1.pdf

Ravenscraft, E. (2022). What is the metaverse, exactly? *Wired*. https://www.wired.com/story/what-is-the-metaverse/

X3d design criteria and specification publication. (2022). https://www.web3d.org/x3d4

X3d online examples and archive. (2022). https://www.web3d.org/example