THE SPATIAL AGE

Evolving the Web for Spatial Computing, Digital Twins, Physical AI, and Robotics



SEAN MANN • DEAN ABRAMSON

A seismic shift has been gaining momentum in spatial computing. As smart glasses become more than just a phone accessory, as digital twins map out factories and entire cities, as physical AI shows up in workplaces and homes, and as robots become untethered from their anchors on factory floors, more and more devices need access to an open infrastructure that mirrors the web while offering intrinsic functionality for spatial computing.

The Spatial Age is the field guide to that shift – from web infrastructure designed for two-dimensional documents to metaverse infrastructure designed for three-dimensional spaces that connects the digital world to the physical world. It cuts through deeply ingrained misconceptions about the metaverse, detailing the convergence of spatial computing, digital twins, physical AI, and robotics, why the web we use now cannot carry that load, and what is needed to build one that can. It illuminates the pathway forward by addressing questions and providing solutions for the critical problems that have stalled the industry.

The only viable path forward is blocked by a single choice being made right now by people in companies who are either not fully aware of, or are willfully ignoring the costly lessons learned from the early days of the web. Sean Mann and Dean Abramson, who are building this foundation in the open, make the clear-eyed commercial case that the open metaverse is not just a noble sacrifice, but the far bigger prize.

armada benson

ISBN-10: 0-9798661-9-7
ISBN-13: 978-0-9798661-9-7
US$ 29.99
52999>
EAN
9 780979 866197

Advance Praise for The Spatial Age

*"**The Spatial Age** exposes a decade of lost potential caused by siloed ecosystems and lays out the roadmap for a future world where physical reality is enhanced by a shared digital overlay – delivered seamlessly from your phone to your augmented reality glasses. Real-world geography governs digital space, replacing physical clutter with clean, intuitive information right when you need it. This is spatial computing at its best.*

To avoid consumer backlash, we cannot rely on a single company. To unlock the true promise of spatial computing, robotics, and physical AI, we must build together through open standards – just as we did at the dawn of the World Wide Web. That vital mission is why the Center for eXtended Reality founded the Open Metaverse Academic Alliance at the University of Rochester.

Anyone building in spatial computing, robotics, or physical AI needs to read this book before deciding how to build our future."

— Barry Silverstein
Director, Center for eXtended Reality, University of Rochester
Former CTO of Optics and Display, Meta Reality Labs

*"Open standards ensure that location information is shared consistently and correctly, and that work now reaches spatial computing and physical AI. People in smart glasses, AI agents, and robots all need to know precisely where they are and what is around them, in a form any device can use. **The Spatial Age** makes a clear case for evolving the web browser so that location, geospatial data, and digital twins can be shared across companies and borders, using open standards that level the playing field."*

— Peter Rabley
CEO, Open Geospatial Consortium

"Before joining XRA, I worked for decades at the Telecommunications Industry Association helping companies agree on the standards that make their products work together. Without these standards mobile communications would never have become what it is today. Most people never see that work, but every phone call, video chat and network connection depends on it. The spatial era is upon us, and the standards to underpin its evolution are still needed. ***The Spatial Age*** *explains clearly why spatial computing, physical AI and robotics now need the same kind of agreement, and how the web can evolve to provide it. This book is a must-read for anyone deciding how their organization will adopt these technologies."*

— Stephanie Montgomery
CEO, XR Association

"Open AR Cloud was founded in 2018 to make real-world spatial computing as open and interoperable as the open web platform, so that anyone could publish content to a physical place and anyone could reach it from any device. ***The Spatial Age*** *makes the same case and shows how evolving the web browser can be key to such a future. OARC is excited to see Sneeze, the open-source engine behind the book, using OGC GeoPose to connect digital content and XR devices to real places. By providing this essential connection between bits and atoms, GeoPose is perhaps the most fundamental interoperability mechanism for real-world spatial computing, with a role similar to what the URL has for the web. We are excited to see* ***The Spatial Age*** *promoting open standards and interoperability as the only viable way of realizing the full potential of the spatial computing future."*

— Jan-Erik Vinje
Co-President, Open AR Cloud Association
Co-Chair, OGC GeoPose Standards Working Group

"Healthcare can't afford a future where every headset, robot and hospital system speaks its own language. ***The Spatial Age*** *makes a clear case for the open standards that will let spatial computing and physical AI work together safely, at scale."*

— Robert Fine
Founder and Executive Director, IVRHA

"Smart glasses will help us reach our full potential by expanding what we can see, hear, and do. That's why we believe smart glasses need an open foundation, so users are free to control their own experience of reality. At Mentra, our enterprise deployments already reveal why context from the physical world matters. AI on glasses can help people get the job done better if it knows which production line they're standing beside, which tool cabinet holds the part they need, or which machine is down. AI smart glasses will bring intelligence into every place we work and live. ***The Spatial Age*** *shows us how we can accomplish this by extending the Web into the world around us. Let's build it together."*

— Cayden Pierce
CEO and Co-founder, Mentra

"For nearly sixty years, I have worked to make virtual and augmented reality serve people by expanding our ability to learn, create, and connect. ***The Spatial Age*** *makes a compelling case for evolving the browser into an open foundation for spatial computing, AI, and robotics. Such a foundation can give more people the freedom to create, share, and benefit from these technologies. This book invites us to build that future together, with the lives we enrich as our measure of success."*

— Tom Furness
virtual reality pioneer
Founder of the Virtual World Society,
the Luminara Institute of Light,
and the family of Human Interface Technology Laboratories

"The internet started closed. AOL, CompuServe and Prodigy were each walled off from the others, and the web only took off once open standards let any device reach any website. Spatial computing and physical AI are at the same starting point today. ***The Spatial Age*** *shows how evolving the web browser can give people, AI agents and robots one open layer for understanding and acting in the physical world. Because no company owns that layer, the data stays with the people and businesses it belongs to, and the world can begin adopting these technologies at scale."*

— Alvin Wang Graylin
Co-author of Our Next Reality
Digital Fellow at Stanford HAI
former President at HTC

"After years of watching as closed platforms have relentlessly suffocated the XR industry, ***The Spatial Age*** *is a breath of fresh air, an inspiring and compelling argument for the necessity of open systems for the emergence of the spatial internet, just as they were for the emergence of the traditional internet before it."*

— Tyler Hatch
Founder and Chairman, The Outernet Foundation

"We build immersive training from our customers' own procedures and equipment, and that content should not be locked to a single device or platform. ***The Spatial Age*** *makes a clear case for an open, standards-based spatial web, and it's the direction our industry needs to go. That's why I'm glad to see the Metaverse Standards Forum doing this work."*

— John Yovanovich
Chief Innovation Officer, CloudIngest

THE SPATIAL AGE

Evolving the Web for Spatial Computing, Digital Twins, Physical AI, and Robotics

Sean Mann • Dean Abramson

armada
benson



Armada-Benson Publishing
Los Angeles, CA
armada-benson.com

Second Edition
Copyright © 2026 by Metaversal Corporation. All rights reserved.
rp1.com

Paperback ISBN-10: 0-9798661-9-7
Paperback ISBN-13: 978-0-9798661-9-7

Electronic ISBN-10: 0-9798661-8-9
Electronic ISBN-13: 978-0-9798661-8-0

Library of Congress Control Number: 2026923913

Printed by Mike Bartholomew / Immersa Media Group
immersamedia.com

Edited by Dean Abramson
Book Design by Dean Abramson
Cover Design by Dean Abramson & Claude Opus 4.8
Illustrations by Dean Abramson & Claude Opus 5.0

All product and company names mentioned in this book are trademarks, service marks, or registered trademarks of their respective owners. Their use does not imply any affiliation with or endorsement by their owners.

Authors' Note
We wrote and published this book with deliberate urgency because the decisions that will determine whether the next version of the web is open or closed are being made right now, often by people who don't realize they're making them. Waiting for a perfect book would have meant publishing after those decisions were settled. We chose timeliness over polish. We expect readers will find places where we could be sharper, fuller, or more accurate. If you do, leave us a note on our website. We'll address it in future editions.

Special thank you to Yin-Chien "Yinch" Yeap for all his contributions.

rp1.com/book/the-spatial-age



Contents

Foreword vii

Preface xiii

Acknowledgements xvii

Part 1 – The Metaverse Extends Beyond Virtual Worlds

1 A Brief History of the Metaverse 3

2 The Metaverse in a Nutshell 11

3 Defining Spatial Computing Terms 21

4 AI, Robots, and Smart Glasses Are Converging 27

Part 2 – Spatial Computing Is Coming of Age

5 An Open Foundation Could Unlock Smart Glasses 37

6 Co-Presence Could Replace Calls and Trips 45

7 Spatial Services Stream In Around You 53

8 Physical AI Delivers Spatial Intelligence 61

Part 3 – The Web and Walled Gardens Both Fall Short

9 It's a Web Browser Problem, Not a WebXR Problem 71

10 Opening Up Is the Bigger Prize 81

11 Enterprises Need Infrastructure They Control 87

Part 4 – Spatial Computing Needs a New Foundation

12 Metaverse Browsers Are Still Browsers 97
13 A Shared Engine Enables the Metaverse..... 105
14 Metaverse Browsers Make Spatial Publishing Easy 111
15 Spatial Fabrics Are the Websites of the Metaverse..... 117
16 The WordPress Moment for Spatial Computing 125
17 Spatial Fabrics Should Control Their Positioning..... 133
18 Each Person Should Own Their Identity 141

Part 5 – Everyone Helps Build the Open Metaverse

19 RP1 Poured the Foundation for the Metaverse 153
20 Together, We Can Build the Metaverse..... 161

Endnotes..... 173
About the Authors 185



Foreword

by Neil Trevett

I have been lucky enough to spend most of my career in rooms full of people much smarter than me, helping competitors work together on standards that grow markets far larger than any one company could build alone. It is fascinating work, though rarely glamorous, and progress is often measured in years. But every so often the pieces click into place, and the industry is never the same again.

Over my 25 years at Khronos, I have had the privilege of helping bring 3D graphics and spatial computing to new platforms. Building on the foundations of OpenGL, OpenGL ES carried hardware-accelerated graphics onto mobile phones. WebGL extended that capability to browsers. glTF gave 3D content a common way to travel between tools and applications, while OpenXR enabled applications to reach devices from multiple vendors. These experiences taught me that, with the right timing and careful nurturing, a standard can create opportunities far beyond those its creators imagined.

When the metaverse returned to the headlines in 2021, Khronos already had many relevant standards. But a globally connected spatial platform would need much more interoperability than we could provide alone. It would take a constellation of standards from dozens of groups, all fitting together seamlessly. Yet there was no common

place to coordinate this work and no shared view of how the pieces should connect. That is why we launched the Metaverse Standards Forum. Its purpose was not to write standards, but to bring the people developing them together to cooperate and communicate. Some 2,500 companies and organizations joined us.

In hindsight, though, much of our industry started down the wrong path, giving the *M-word* a bad rap. The focus was on virtual worlds, and the metaverse was widely pictured as a place you escaped to. Strap on a headset, shut out the real world, and live inside a simulated one. For many, that vision felt isolating at best and dystopian at worst, and I don't think they were wrong. Perhaps that should not have been a surprise. The word itself was coined in Neal Stephenson's Snow Crash, a novel set in a bleak dystopian future.

Meanwhile, the industry was splintering into a series of walled gardens, offering little prospect of the openness that had made the 2D web so successful. And company after company seemed surprised when their own siloed version of the "metaverse" foundered.

Then AI came into the mix and began to change what spatial computing could be. Scene understanding and natural interfaces are making the world machine readable. Smart machines and people wearing emerging devices such as smart glasses are beginning to engage with the real world itself in ways that bring people together rather than shut them away. The metaverse could now deliver on its real promise of connecting people, AI agents, and devices through spatial computing. Perhaps we could rehabilitate the *M-word* after all.

Most of the pieces of the technology jigsaw were on the table, but I still hadn't seen the picture on the box. That is when I met Sean, Dean, and the team at RP1.

Their central idea was deceptively simple, though its implications took time to sink in. The metaverse should be a vast collection of spatial services, provided by many independent parties and composed live around whoever or whatever happens to be in a given place. Imagine a maintenance engineer looking at a machine through smart glasses. The manufacturer might supply repair instructions, the operator could provide current operating data, and an independent service agency could connect the engineer with a remote expert. The

value comes from those services working together in context, even though they originate with different providers.

It took some patient explaining on their part, but once I saw the metaverse in those terms, so much suddenly made sense. Of course it has to be services. Of course they have to come from many sources. And of course no single company can own the composition. Here, at last, was a path to a global, standards-based platform, one where anyone could easily build and deploy their own spatial content. That openness is what made the web pervasive for 2D content. Now it was time for the web to step off the screen and into the world around us. This is the way.

But with standards, a vision alone is never enough. It takes someone with a real stake in the outcome, and RP1 could see that an open metaverse would create a market where their spatial technology could add significant value. Enlightened self-interest is the fuel of many successful standards. They also had the skill and determination to make it work. Sneeze + Rubidium, the open-source engine and browser at the heart of this book, and Artemis, their native metaverse browser built on it, are real, running software, not slideware. And they have had the tenacity to explain the vision, again and again, patiently bringing others along with them, me included.

As the pivotal role that standards would play became clearer, RP1's vision and the Metaverse Standards Forum turned out to be an ideal match. The Forum offered neutral ground for weaving the interoperability threads, and a natural home for an open-source browser engine that no single company would control. And so the Open Metaverse Browser Initiative, or OMBI, was born. It does not try to replace the existing web. Instead, it develops a metaverse browser engine that builds on today's web and adds the ability to connect to and display spatial services.

The Spatial Age makes this vision concrete, turning it into an architectural proposal readers can examine, question, and help develop. For those building spatial technologies, it offers a way to consider how their work might connect to a larger ecosystem. For readers trying to understand where this field is heading, it explains why the relationships

between services matter as much as the devices through which we experience them.

There is a great deal of work still ahead. But Sean, Dean, and the team have given us a picture of how the pieces fit together, and running software with which to test it. That gives the industry a starting point to build on and improve together. I am honored to be a part of the OMBI journey.

The spatial age is coming. Whether it arrives open, and belongs to everyone, is up to us.

— Neil Trevett
President, The Khronos Group and the Metaverse Standards Forum
Vice President of Developer Ecosystems, NVIDIA



Preface

Somewhere along our journey, building infrastructure for spatial computing with our team at RP1, sitting on the board of the Metaverse Standards Forum, helping launch an open-source browser engine for the physical world, and building the first native metaverse browser on top of that engine, we arrived at the conclusion that we now hold with complete conviction: almost everyone has the metaverse wrong, and the misunderstanding is costing the spatial computing industry a decade.

The entire world has already begun moving from web infrastructure to metaverse infrastructure. A dimensional shift from an internet of flat pages we consume to a spatial layer we navigate in both the real and virtual world. The web we rely on now, built for documents on a screen, was never designed to handle that shift.

Shifting from the web to the metaverse will require evolving not just websites, but the web browser itself. While this may seem like an insurmountable task, we have four decades of hindsight to use as our guide. Nearly all the mistakes the spatial computing industry is making today are ones we've seen in the past. History has handed us the blueprint for finding our way out of the maze of walled gardens and closed ecosystems and into an open, evolved web – better known as the metaverse.

The metaverse is not a game. It is not a place you escape into. It is not virtual reality, though you may sometimes experience it through

a VR headset. The metaverse is something much larger and far more consequential. It is the evolution of the World Wide Web by which the internet transcends the screen and inhabits the physical world around you. It is the web learning to understand places, objects, machines, and people, and it is arriving right now.

Spatial computing is the form this evolution takes – the virtual layered onto physical space. Artificial intelligence will accelerate its adoption at a rate that would have been hard to imagine just a few short years ago, letting our devices, our software, and our data finally perceive, understand, and act in the world around us. These two forces are converging at exactly this moment, and our ever-increasing dependence on them will shape the foundation of the metaverse for decades to come.

Tens of billions of dollars are funding the shift into the metaverse every quarter. Most of the biggest companies on earth are either building it or seeking out ways to operate within it. These companies include robotics labs, physical AI and world-model companies, makers of AR glasses and digital twins, and tool builders giving the world new ways to capture and create in 3D. They also include nearly every Fortune 500 company. Yet few realize that they are all converging on the very same thing – a shared, open layer that connects the digital world to the physical one.

We want to be honest about our vantage point. We are not neutral observers. Our company is building in this space, and we have staked our work on the belief that this foundation must be open, like the web itself, and not owned by any single company. That belief is not charity. It is the only version of this future that actually works, and a large part of this book is spent explaining why.

This endeavor demands a fair amount of humility and patience. Helping to shape open standards, the unglamorous plumbing that replaces today's fragmentation and friction with real interoperability, so that devices, software, services, content, people, and robots can all work together instead of in isolation, is hard, slow, thankless work. Most companies avoid it. They build moats instead, because moats are how you recoup research, development, and operational costs and turn a profit. We understand the instinct. But we have become convinced

that in this particular transition, the moat-builders will fail. Not despite their moats, but because of them. The irony is that the companies spending the most on building their moats are the ones with the most to gain from a future built on the open metaverse.

Until we solve the underlying reasons why enterprises, universities, governments, and ordinary people are both unwilling and unable to adopt these technologies at scale, none of the billions being poured into smart glasses, digital twins, AI models, and robots will pay off. That problem is not a product problem. It is an infrastructure problem. And infrastructure the whole world builds on top of cannot belong to any one company. Like the web before it, the metaverse must be built in the open.

The purpose of this book is to explain where we are headed, why the web we have now cannot take us there, and how we build the one that can. We intend to show you the future vividly enough that you will seriously consider what you can do to help us get there. If we do our job well, you will finish this book unable to unsee what we see – that the metaverse is, and always was, a spatial evolution of the web itself, reaching off the screen and into the physical world.



Acknowledgements

by Sean Mann

Chris Ferguson, *our co-founder* and earliest believer. Your support gave us the freedom to go out and execute with the confidence to change the world. Through every pivot and turn, you never wavered. There would be no RP1 without you. Thank you for your trust in us trying to find the oasis.

David Sernaque, *our master problem solver* and one of the most extraordinary programmers I have ever met. No technical challenge is too hard for you, and this work has given you plenty. You have known Dean longer than any of us. Watching the two of you work together is like watching an old married couple write software – equal parts engineering and argument, followed somehow by the right answer. Thank you for sharing Dean with the rest of us.

Jim Lafferty, *our powerhouse* and jack of all trades. You designed and run the infrastructure everything we build depends on. Every demo, every customer, and every line of code lives on the systems you keep running. You never say no to a new project, no matter how full your plate is, and we know anything we hand you will get done. Listening to you and Dave debate sports and politics is the highlight of our Monday mornings.

Yin-Chien "Yinch" Yeap, *our champion of open standards* and our voice of reason. When I told you we wanted to build the open metaverse, you laughed, then realized I was serious, and somehow you were crazy enough to join us. Thank you for advocating on behalf of every developer and creator who has been burned by one too many walled gardens, and for turning the improbable into the everyday with us. Your assistance with this book brought the same clarity and rigor to these pages that you bring to everything we build.

David Nassau, *our marketing mastermind* and Red Vines connoisseur. You do the work of an entire marketing department, and your guerrilla instincts have put RP1 in front of people we never could have reached otherwise. Your energy and passion for what we are building are contagious, and your ability to befriend so many people is a rare gift, one that makes traveling the world with you an adventure.

Junior Bustamante, *our graphics guru* and creative wizard. You bring our brand and web presence to life. You don't just build webpages; you understand how the web actually works, which is exactly what we need. You take input from every direction with remarkable patience and transform it into something better than we imagined. Thank you for keeping pace with every change of direction, and for making each one look like it was the plan all along.

Dean Abramson, *my friend* and co-founder. You are the architect of the technology that makes all of this possible, and you see things the rest of us can't see yet. The best version of this future sends people back out into the real world rather than deeper into a screen, and you understood that before I did. Without your vision, sacrifice, and relentless drive, none of this would exist. Thank you for the long conversations, the good arguments, the breakfast burritos, and the laughs. There is no one I'd rather build this with.

Neil Trevett, *our biggest advocate*. You have spent your career doing what few can – bringing an entire industry together around shared standards. When I first shared our vision with you, you were curious, but with a healthy dose of skepticism. Once you got it, we all knew, *this is the way*. Your belief meant more than you know. Standards work is too often thankless, so let us say it plainly – thank you for your guidance and for helping bring this work to the standards community.

Everyone in this industry, *who kept building and laying the groundwork,* often long before the world was ready. Thank you. We hope the standards we are creating bring us all together and invite the world to adopt these technologies at scale.

To our families, who never wrote a line of code
but made this all possible…

Rae, *my wife and best friend,* you're the first person I want to tell when something goes right, and the only one I want beside me when it doesn't. You believed in me from the start and stood by me through all of it. Our best chapters are still ahead. I love you.

Logan and Riley, *my children,* you amaze me every day. Being your dad is the best job in the world. Thank you for your support. You've been my inspiration every step of the way. Don't be afraid to fail, some of your best lessons will come from it. Wherever your dreams take you, I'll be cheering the loudest.

Richard and Linda, *my parents,* everything I've built started with what you taught me, *work hard, treat people well, never stop learning.* I miss you terribly and wish you were here for this. I hope I have made you proud.

Joy, Adam, and my twin, Shane, *my siblings,* we share a lifetime of memories, and I am grateful we have had each other through it all.

Lou and Judy, *Dean's parents,* thank you for welcoming me into your family and for supporting Dean and me through every step of this journey. I treasure the meals, the laughter, and our evenings playing Oh Hell, especially the rare nights you let me win. You always make me feel like I have another home.

Part 1

The Metaverse Extends Beyond Virtual Worlds



Chapter 1

A Brief History of the Metaverse

"All I did was make up a name for it, and put it in a novel that got read by a lot of techies."

— Neal Stephenson
author of Snow Crash

Fiction led us to mistake the metaverse for a virtual world

The metaverse was born in fiction. Neal Stephenson invented the term in his 1992 novel, *Snow Crash*, to describe a virtual successor to the internet.[1] Good science fiction paints a picture for us to envision, and this picture stuck. For thirty years people deliberated about what the word meant and where the metaverse might transport us to. Those arguments weren't merely academic. New words need to be vigorously dissected and analyzed before they can be used to communicate a common meaning.

The challenge was, an idea borrowed from fiction required something tangible to attach itself to in order to have real world meaning. The path of least resistance was to attach the concept of the metaverse to the digital worlds that already existed, the ones that most

resembled the fiction. Second Life, which arrived in 2003, was among the earliest examples of this. It allowed people to build and inhabit a persistent digital world, buy virtual land, and attend virtual events. Then came the MMORPG worlds like World of Warcraft, and large online multiplayer platforms like Minecraft, Roblox, and Fortnite, each a massive, persistent, social space where millions of people devoted countless hours and constructed virtual economies bigger than many small countries.

When journalists, analysts, and tech executives reached for a word to describe these futuristic virtual worlds accessible only through computers, laptops, and phones, they found them not too dissimilar from the one described in *Snow Crash*. The association stuck, and they became the de facto real-world definition of the metaverse. By 2022 the RAND Corporation was describing it as "a catch-all term for anything involving a blend of AR, VR, digital avatars, and blockchain-related technologies."[2] When a word means *anything*, it conveys *nothing*.

Mistaking closed platforms for the metaverse repeated AOL's error

The earliest days of the internet provide a clear analogy to see just how big this miss was. Before the web had become commonplace, people were led to believe that AOL was the internet, thanks in no small part to one of the largest direct mail marketing campaigns in history. Its CEO, Steve Case, went so far as to declare, "It's not a choice between AOL and the internet… AOL is the internet and a whole lot more."[3] To the contrary, AOL was a curated, walled-garden, privately-owned service that did not interface with any of the open standards and protocols of the web until consumer demand forced the company to do so – albeit from within the confines of their closed system. That fact still did not sway most people from equating AOL with the internet or the World Wide Web at the time. The world would soon discover that AOL was actually the antithesis of the web.

That is, more or less, what happened to the metaverse. One small, closed proxy of it mistaken for the whole. Even as late as the early years of this decade, some of the most respected thought leaders in the industry continued to reinforce the wrong picture. In one of the most

widely quoted definitions, Matthew Ball dubbed the metaverse, "a massively scaled and interoperable network of real-time rendered 3D virtual worlds."[4] At best, this definition is entirely misleading. Calling the metaverse a set of 3D virtual worlds not only perpetuated the incorrect perception that the metaverse was nothing more than a collection of proprietary, walled-garden systems, but it entirely fails to recognize any link between the digital world and the real world – one of the metaverse's most valuable and defining characteristics.

Meta's rebrand made the narrow definition dominant

In October 2021, the pervasive, but misguided belief that the metaverse was a collection of 3D virtual worlds led directly to a trillion-dollar corporate bet. Facebook rebranded itself as Meta and staked its entire future on this narrow interpretation, pouring around one hundred billion dollars into the development of their own, closed, 3D virtual world and a line of virtual-reality headsets for consumers to access it.[5] The public heard one message loud and clear. The most valuable social-media company on the planet had just declared that the metaverse was: you, in a headset, as an avatar, in a virtual room. Their marketing was unambiguous. The metaverse equals a virtual reality 3D world accessed exclusively via an uncomfortable VR headset.

For a moment, the hype was everywhere. Consulting firms rushed out metaverse reports. Brands tinkered with virtual storefronts. Then people tried it, and the backlash was unrelenting.

The collapse of consumer VR turned the metaverse into a punchline

By 2023, the metaverse had become a punchline. The virtual worlds were sparse and awkward. The headsets were heavy and isolating. The promised crowds did not materialize. Staggering losses were reported.[6] Fashionable opinion pivoted from breathless to disdainful almost overnight. The *M-word* had become so toxic that companies working on genuinely important spatial technology began scrubbing every last mention of it from their materials.

The clunky VR anti-social worlds that were being endlessly mocked genuinely deserved a hefty dose of skepticism. The safe move,

the smart move, was to steer clear of any reference to the metaverse in corporate settings. It was an unfortunate consequence that the mockery was so tightly bound with the term *metaverse* itself. When a more authentic and applicable definition of the metaverse began to arrive, the executives, investors, and journalists who were on the front lines of recognizing innovation had already written it off.

The misunderstanding diverted institutions, investment, and innovation

The mischaracterization of the term *metaverse* was not without consequence. It changed where money was allocated, where talent was assigned, and ultimately what was produced. It did three quantifiable kinds of damage.

First, it told serious institutions such as hospitals, airports, factories, universities, and government agencies to look away. These organizations, whose adoption determines whether a platform shift succeeds, heard *metaverse* and reasonably concluded it was not for them.

Second, for those who were not scared off, it aimed capital and talent at entertainment instead of infrastructure. Perceived as little more than virtual worlds, metaverse content was constrained to 3D scenes, avatars, cosmetics, and social experiences. The foundational work that underpins the metaverse – how devices find and connect to spatial content, how identity travels, and how independent services compose into one scene – received almost no attention.

Third, and most damaging, it stifled innovation by restricting development to a tiny group of very wealthy companies who could afford to build and promote these extremely expensive endeavors. Future profits from monopolistic (or duopolistic) app stores justified the impetus and business model for these ecosystems. But they can only work within the confines of tightly controlled, closed, proprietary, walled-garden platforms.

Recasting the metaverse as an open system reveals the real transformation

Given the mountain of evidence in front of them – awkward headsets, empty virtual rooms, and a parade of bankrupt companies –

it's hard to argue with the skeptics. They certainly were not wrong about the fact that a vast majority of adults are not interested in spending their free time being harassed by adolescent kids whose behavior and foul language would make a drunken sailor blush.

In hindsight, we can probably agree that never was the transformation the world was hoping for when the picture of the metaverse was first presented. But, right around this same time, a small butterfly in the form of a novel idea crossed paths with the metaverse, introducing a subtle change to one of its underlying assumptions. This one tiny idea, if adopted universally, would have profound effects on how the metaverse fundamentally operates.

What if the metaverse was never meant to be a collection of closed systems, but rather one single open system – an evolution of the web itself?

Then walled gardens are exactly the wrong architecture for the metaverse. The web proved this case more than three decades ago. An open metaverse that extends the web – one based on open standards, open protocols, decentralized infrastructure, and browsers that operate on every device – could easily become the basis of a generational transformation the corporate world has desperately been in search of. Satya Nadella, the CEO of Microsoft, said, "The metaverse enables us to embed computing into the real world, and to embed the real world into computing, bringing real presence to any digital space."

While the pundits were busy burying the *M-word*, the transformation of the web into the metaverse was arriving, not as an escape from the real world, but as the real world coming online. This evolution from two-dimensional information to three-dimensional spatial context had begun to gather momentum in robotics labs, by AI models learning to see, within the smart-glasses divisions of the world's biggest companies, and in the digital twins quietly running inside factories and hospitals. All of these components more closely resembled the products and services found on the open web. The problem, still, was that these spatial technologies were all being developed separately using ad hoc technologies originally designed for 2D infrastructure.

Reclaiming the term restores its original promise

Even though this new vision of the metaverse has gained traction, the wariness surrounding the term *metaverse* lingers, running especially deep in Europe and in enterprise boardrooms, where the consumer-VR hype landed with a spectacular thud. Given all the history behind it, a fair question is whether the word is worth keeping at all. If it is so bruised, why not just coin something clean and walk away from the wreckage?

Several have tried, quietly swapping the *M-word* for safer labels such as *spatial web*. It turns out that getting the masses to agree on a new term's meaning is easier said than done. Every fresh label adds yet another moniker to an industry that already has more than its fair share.

Pronouncing the metaverse dead may have been fashionable for a while, but it now appears to have been a bit premature. The people building the industrial version of this future have never stopped using it, though they have continued to evolve and refine its meaning over time. Today, the metaverse is most accurately and succinctly described as an evolution of the web, extended into the physical world.

Standing in front of shareholders in 2024, Roland Busch, the chief executive of Siemens, was asked, "What exactly is the industrial metaverse?" Without hesitation he replied, "It's a place for industrial innovations. It combines the real and the digital worlds. A virtual, photorealistic space where people – and AI – can work together in real time to solve real problems."[7] Across manufacturing, energy, and logistics, the early adopters keep reaching for this same word, because it is perfectly suited for, and the only one expansive enough to describe what they are building.

Therefore, we are openly and unabashedly reclaiming the term *metaverse* rather than renaming it. A word with baggage, sure. But there is no better alternative. Besides, the problem was never the word itself. It's been a terrific word ever since its inception in a 1992 novel, complete with a picture that isn't all that far off from what is actually happening today. All it required was a simple idea to flutter by – the notion that it needs to be founded on a single, world-wide open network – and the word turns out to be exactly right. The metaverse is the evolution of the web, extended into the physical world. It is the

foundation for spatial computing, digital twins, physical AI, robotics, and all the ways we will interact with them in both the real world and in virtual ones alike.



Chapter 2

The Metaverse in a Nutshell

> ***"We are all now connected by the Internet, like neurons in a giant brain."***
>
> **— Stephen Hawking**
> *astrophysicist and author*

The metaverse extends the web from documents to places

The web connects people to documents and information. Any person, using any device, can reach any page that anyone else chooses to publish and without needing permission to do so. This revolutionary idea that was spawned four decades ago grew into the single most significant innovation in the history of humankind and continues to provide the world with open and universal access to unlimited amounts of information.

Roughly 500 exabytes (or 500 million terabytes) of new data were created worldwide every day in 2025.[8] The overwhelming majority of this data is two-dimensional. Think text, documents, images, and videos. However, with the increasing quantity of spatial devices and services that operate in 3D space such as smart glasses, robots,

autonomous vehicles, and digital twins, more and more of that data will require new, unique ways to interact with it spatially. At best, some of these new paradigms are ill-suited for use in web browsers and mobile devices. At worst, they are treated as hostile invaders to a web browser that was architected to handle documents from one publisher at a time.

The metaverse solves these issues by applying the same characteristics of the web to places instead of documents, but with fresh architecture that is more adept at handling spatial data. Rather than typing a web address to reach a page, all relevant information and services that occupy a location like a busy street intersection are automatically activated as you walk through it, each anchored to the buildings, stores, and objects in front of you. Like the web, the content and services you interact with will come from many independent providers. However, unlike the web, they will need to be presented in one unified scene, all contributing to what you see and experience concurrently and contextually in 3D space, rather than in isolated 2D windows.

Attaching digital information to real places is not nearly as radical as it sounds. Large parts of the physical world have already been mapped into virtual counterparts – digital twins. This includes entire cities like Singapore, Dubai, and Tokyo, just to name a few. Mapping software like MapQuest and Google Earth have been around since the early days of the web. And every day, companies bring their factories and warehouses to life in one digital form or another. BMW's plant in Regensburg, for instance, runs a complete virtual replica of its factory floor, detailed enough that production changes are tested in simulation before a single physical part is moved.[9]

Furthermore, digital content anchored to real places is not new, as anyone who has carried a phone for the last ten years can attest. Google Maps has been drawing routes onto streets for years. Pokémon Go put characters on real sidewalks a decade ago. Snapchat sticks objects to real faces and real rooms, and your phone will redraw a sign written in a language you do not speak. All that works today.

The metaverse is multi-origin, spatial, and driven by proximity

Among a handful of differences between the web and the metaverse, one particular feature stands out. The web is intrinsically architected from the ground up to be single-origin – meaning, each website sits in a silo of its own. From the ground up, the core of every web browser is designed to prevent any website (or origin) from encroaching on another. This assumption is so tightly ingrained into web browsers that two separate hosts from one domain must jump through complicated hoops to share information between them, while some rudimentary and necessary features (like storage) simply don't allow it at all.

The single-origin architecture affects how you search in every way. You visit the web one page or one app at a time. Two different pages or apps have no spatial relationship with each other. Unlike the real world, bedbathandbeyond.com does not sit adjacent to starbucks.com like they would in a strip mall. You decide to load each website as you need, not because they're near you. Independent webpages don't overlap or intertwine. Each one sits neatly inside its very own, and very isolated rectangle.

This is not inherently bad at all. In fact, it has substantial benefits. This model of browsing has served us exceptionally well because we process two-dimensional information very efficiently this way, and because the form factor of the devices we mainly use today are primarily designed to arrange 2D content into rectangles. Web browsers and the devices they run on, however, are not very well suited for working with spatial 3D content from many origins at once based on where you happen to be standing right now.

As luck would have it, metaverse browsers and the spatial devices they run on are exceptionally well suited to solve exactly these challenges. Unlike with the web, the very nature of the metaverse is designed from the ground up to be multi-origin, meaning lots of services all share a common scene equally, none being more important than any other. Furthermore, those origins are all loaded based on their proximity to your location, not because you typed a URL or clicked a hyperlink.

Web browsers naturally provide intrinsics for working with 2D information, such as the ability to display and layout text, graphics, images, and videos. But they lack most of the intrinsics needed to work in a spatial context. That's where metaverse browsers pick up the slack. They provide built-in capabilities, which include GPS and VPS for locating where you are in the world, identity, persona, and avatars for co-presence, world mapping, and 3D immersive graphics, among other intrinsic spatial features.

The metaverse preserves the web's openness, self-publishing, and owner control

Despite its differences, because the metaverse is an evolution of the web, many of the vital aspects that make the web so powerful remain unchanged. The most important among them, the metaverse is a single open system. That's because it shares the exact same internet infrastructure – we don't need a separate internet to run the metaverse. This attribute is the overarching reason why so many other underlying features are preserved, all of which are essential for enterprises who see these features as non-negotiable for their business operations.

A company that wants digital content to appear over its own property can freely publish that content on its own servers, in precisely the same way that it operates its own website. No other company sits in the middle of that arrangement, which means nothing is submitted for approval before it goes live, nobody takes a share of the revenue it earns, and customer interactions and data remain private. The equipment and the skillset required to create and publish content in the metaverse are the same ones already needed to run a website, so anyone who knows how to publish on the web can do so on the metaverse.

An individual or company must retain the right to operate the space that covers their home or business without fear that someone else can publish content that would interrupt their business. That doesn't preclude others from bringing items into those spaces or overlaying content on top of them. Anyone can turn on optional layers that sit on top of the virtual world – like augmented, augmented reality. But each person must choose to turn on those layers; they cannot be forced on anyone.

A Yelp overlay would be a perfect example of how this operates in practice. A restaurant or barber shop would run their own metaverse location as needed to operate their business. Yelp would add a layer on top of that for users to see reviews of those businesses. But Yelp could not force you to see their reviews if you didn't want them. This is no different than the way most civil societies operate today. A competitor cannot put a sign in another company's place of business. Nor can people arbitrarily put up their own traffic signs.

No device, app, game, or virtual world is itself the metaverse

At this juncture, it would help to clear up a few misconceptions about the metaverse. Here are four things that are routinely mistaken for the metaverse, but that are definitively not.

A VR headset is not the metaverse. It's one of many pieces of equipment for viewing the metaverse, though. A monitor is the analogous piece of equipment for looking at the web. Very few people would say that a monitor is the web.

A proprietary, three-dimensional virtual world published by one company is not the metaverse. In fact, such worlds are the antithesis of the metaverse. The metaverse is not defined by how many people can be co-present in a given space. Only that it participates in the open spatial protocols that guarantee the ability to self-publish content and directly communicate with a metaverse browser. A platform that requires proprietary software to be downloaded and installed, that issues its own user accounts, personas, and avatars, and that does not interoperate with other metaverse spatial content or services could hardly be the metaverse.

A game is not the metaverse. The metaverse will definitely have games, in the same way that there are games on the web. But games account for only a teensy tiny fraction of what the web is used for. Games will likely account for a similarly small sliver of what the metaverse is used for.

A phone application that attaches virtual content to a real location is not the metaverse. It is merely one company performing one service. The metaverse will be made up of millions of these services – or even

hundreds of millions – in the same way the web is made up of hundreds of millions of websites.

Roblox and Fortnite are both perfect examples of things mistakenly described as metaverses because: for one, there is only one metaverse in the same way that there is only one web; and for two, they operate entirely on proprietary protocols and principles that share nothing in common with an open metaverse. Whereas Facebook, a website, is clearly a participant of the web, Roblox and Fortnite can't even claim to be a participant of the metaverse given that they are both entirely closed platforms.

The metaverse makes physical, virtual, and remote places equally reachable

The metaverse can be accessed with equipment you already own. Desktop computers, phones, and tablets provide adequate access in most cases. However, it is probably best experienced with smart glasses and VR headsets. These two devices each have a particular strength when it comes to working in the metaverse.

Smart glasses add to what you can already see, holding digital information in fixed positions in the space around you, leaving your hands free and your eyes up. That is augmented reality, or AR. A VR headset, on the other hand, replaces your view entirely, putting you in a distant location far more convincingly than any flat screen can. That is virtual reality, or VR.

Whereas AR can only be used to experience the world in which you are currently present, VR has two distinct purposes. First, you can use a VR headset to teleport to another location on the planet – provided that location has a metaverse digital twin to travel to. While accessing a remote location with your VR headset, you can interact equally with the AR users who are physically present at that location and the VR users teleporting in from elsewhere. Second, you can use VR to travel to an entirely imaginary location that only has a virtual representation in the metaverse. While there, you can interact only with other VR users, of course, since no one can be present in a virtual space with AR glasses.

The metaverse requires an order of magnitude more from the internet than the web

The future metaverse will need to process way more data than the web handles right now. Approximately 217 million active websites exist today.[10] Compare that with roughly 2 billion internet-connected buildings on earth.[11] That is nearly ten times as many buildings as there are working websites. And that doesn't even account for all the connective areas between them such as streets, highways, parks, and plazas that municipalities own and maintain. Those spaces already carry information in the form of signs, markings, and signals, but will contain abundantly more active services and data once their digital twins get produced.

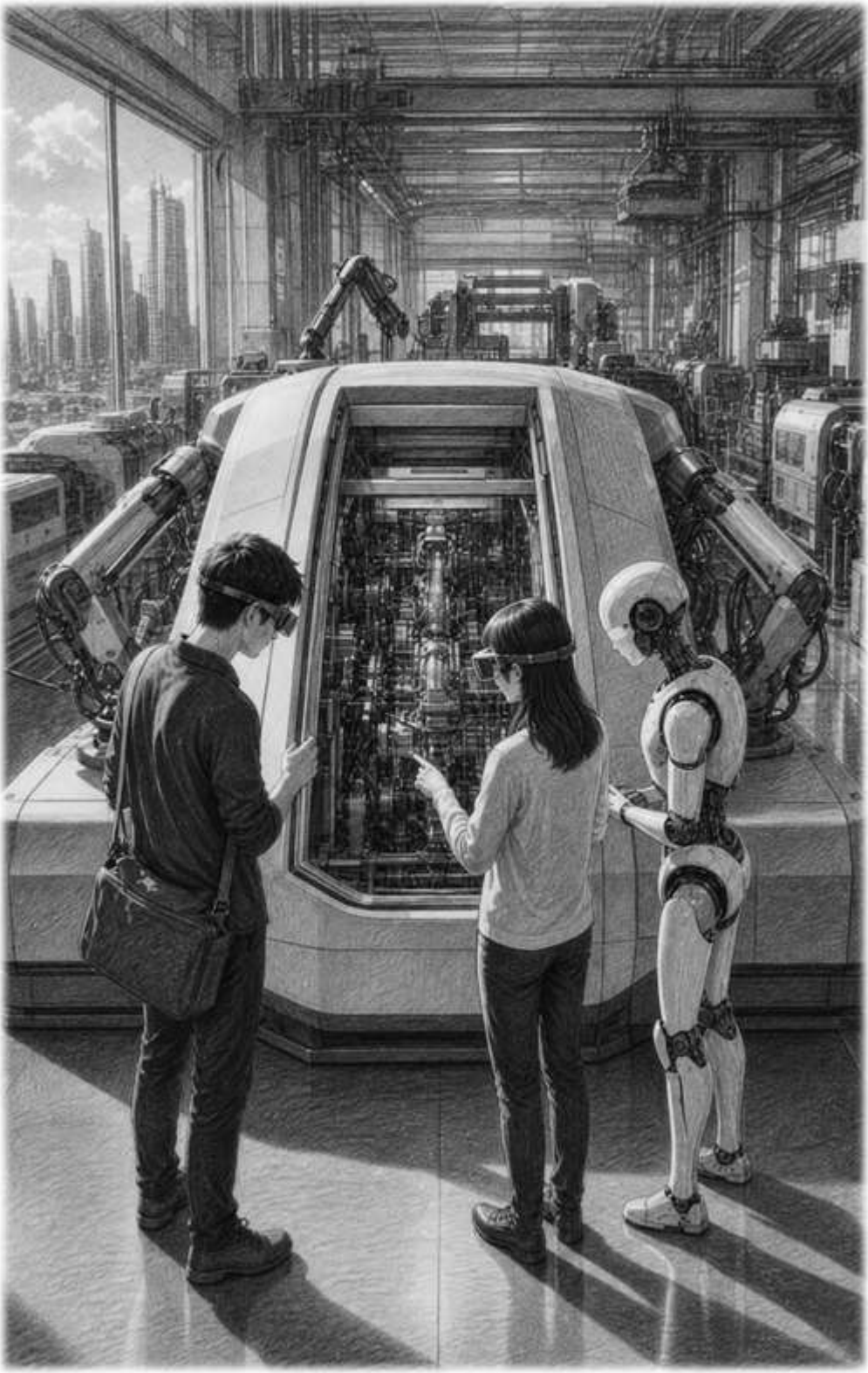
The nature of both the frequency and quantity of data transmitted will increase dramatically as well. A typical web connection spends most of its life doing nothing. A webpage request transmits a burst of data at first, but then it sits idle while you read it. While video accounts for about two-thirds of all internet traffic,[12] video page requests account for less than 1 percent of total page requests. While the math is admittedly fuzzy due to the number of variables involved, if 100 percent of metaverse users consume one-fourth the bandwidth of a video, that points to an order of magnitude increase in total internet traffic.

The three biggest streaming services (Netflix, Disney+, and Amazon Prime) plus the two biggest video sites (YouTube and TikTok) combined account for more than a third of all internet traffic. Even though that is a massive amount of data, the prerecorded streams you watch are cached on edge servers as close as possible to you. The metaverse demands a completely different type of traffic. Real-time co-presence and service-related data are not cacheable at all and must be transmitted instantaneously from all corners of the globe. It is unknown if the current internet infrastructure can support an order of magnitude increase in traffic that is all real-time.

Taken together, the demands placed on the internet by metaverse infrastructure will require ingenuity to solve. This includes changes to both the server and client software systems necessary to support the forthcoming evolution of the web into the metaverse. The effort

involved will also require participation from a wide spectrum of technology leaders that exist today, as well as several new ones that will undoubtedly spring up to solve new problems as they are discovered.

This work is underway, led by a small handful of people and companies who are setting the standards of the open metaverse and producing open source versions of metaverse browser engines and metaverse browsers. Other companies, universities, and government agencies are contributing by working on tools and digital twins that follow these standards. The advent of AI will accelerate this process exponentially. No one – who has read this book – should be surprised when this happens within only a few short years.



Chapter 3

Defining Spatial Computing Terms

"If you wish to converse with me, define your terms."

— Voltaire

writer, philosopher, and historian

XR describes the experience, with AR overlaying reality and VR replacing it

A lot of confusion in boardrooms and headlines comes from people using the same word to mean different things. To avoid confusion for the remainder of this book, here is a set of concise definitions needed so we are all speaking the same language.

XR is the umbrella term for immersive augmented or virtual reality experiences. Per the Khronos Group, the originators of the term, the letter X acts as a symbolic placeholder, and specifically is not an initial that represents the word *eXtended*. You experience virtual content added to your field of view or wrapped fully around you through an XR device, but XR refers to the experience itself, not the

hardware that delivers it. XR is the content you experience within the metaverse, the same way webpages are the content you view on the web. Under that umbrella sit two main flavors.

AR (Augmented Reality) overlays digital content onto the visible real world in which you are present, comprising most of your view. This can occur in many forms from simple graphical overlays such as arrows pointing you towards your destination to interactive objects that are aware of your physical surroundings like a virtual cat that knocks a virtual vase off your very real counter (because that's what virtual cats do). Today this already works on mobile phones and tablets, among other devices. But a phone can only show you the narrow window of the world that fits inside its relatively small display. AR glasses will one day open that up to your whole field of view, which is why glasses, not phones, are expected to be the preeminent device most people will use to interface with the metaverse. These devices can also be labeled as smart glasses when paired with AI.

VR (Virtual Reality) replaces your view of the world with a fully rendered one, either real or imagined, in which you are not present in real life. This is typically achieved by wearing a somewhat bulky VR headset that provides separately rendered images for each eye combined with stereo audio. This provides an immersive experience that enables you to be present in places other than where you currently are in real life such as a factory on another continent or an intergalactic cruiser making its way to the Alderaan system. Today's bulky headsets are temporary. VR hardware is shrinking year over year and the latest headset from Meta almost resembles a pair of glasses.

MR (Mixed Reality) is a term that refers to some combination of augmented and virtual technologies. To add to the confusion, Microsoft adopted the term to differentiate their defunct HoloLens product from other AR products on the market despite exhibiting little or no difference from them. While there may or may not be a technical dividing line between AR and MR, most people in the industry honestly could not tell you what that line is and simply use the more common AR term. We will not use MR in this book.

Spatial computing lets devices sense and operate in three dimensions

Spatial computing is computing that senses and operates in three-dimensional space. A spatial device uses sensors like cameras or GPS to know where it is, perceive the geometry of its surroundings, and can place digital content at precise physical locations that remain fixed, even as you move. Where traditional computing presents two-dimensional content on computer screens unrelated to any physical location, spatial computing deals in places, surfaces, and objects, each assigned a position within a coordinate space. AR glasses and VR headsets are displays for spatial computing, which enables them to anchor a maintenance alert precisely above a valve and keep it in place as you walk around the machine. Smart glasses, phones, robots, and vehicles are all spatial computers.

Physical AI perceives and acts in the real world

Physical AI is artificial intelligence that perceives and acts in the physical world, rather than living entirely in conversation windows. A large language model that writes an email is not physical AI. A model that lets a robot see a cluttered shelf, decide how to grip an unfamiliar object, and carry it across a room is. Physical AI is what turns robots, drones, and autonomous vehicles from pre-programmed machines into adaptable agents that can operate in messy, unscripted human spaces.

Digital twins mirror reality while world models learn its rules

Digital twins are live, data-connected virtual replicas of something in the real world, whether a machine, a building, or an entire city. They are not just static 3D models. They stay synchronized with their physical counterparts through sensors and update as reality changes. Digital twins are used to monitor, simulate, and control their real counterparts using an array of sensors and components that are built into or surround them.

World models are AI systems that learn to understand and generate accurate three-dimensional objects and locations from text, photos, or video. They are also capable of modeling geometry, physics, and

dynamics. They can turn photos or videos of a place into an explorable 3D environment or generate new spaces that obey the rules of the physical world. Where language models learn the patterns of words, world models learn the patterns of worlds: how light falls on a surface, how objects rest and collide, how a space holds together as you move through it.

Visual positioning locates devices and content precisely

VPS (Visual Positioning System) uses GPS combined with a device's camera and a pre-scanned map to determine precisely where you are in the real world. It is far more accurate than GPS, and it works indoors, where GPS fails. Without VPS, it wouldn't be possible to anchor virtual objects to the physical world and keep them steady as you move around a location. For instance, a remote colleague could meet with you on his VR device in your office, but your AR glasses wouldn't be able to properly position him in the chair beside your desk unless your glasses knew exactly where your chair is relative to you and your room.

Identity proves who you are, personas express context, and avatars embody them

Identity is who you are as a human being, along with the protocols that let you prove it. Done correctly, it lets you authenticate yourself in the metaverse without needing to reveal personal details about yourself to whoever happens to be asking. Your identity wallet can prove individual facts, like being over eighteen or licensed to do a job, without revealing anything more about you, including your real name.

Personas are how you choose to publicly present your identity in a given place. You probably have more than one personality in the real world – the version of you that shows up at work is not the same as the one that plays games with friends. Likewise, a single identity can hold many personas, each suited to a different occasion in the metaverse. However, it is vital that none of your personas be traceable back to your identity, nor to each other.

Avatars are the visible form that each of your personas take in the metaverse. They are how you show up when you cannot be somewhere

in person; a representation of you placed into a shared space so others can see, hear, and interact with you. Your avatar might be a photorealistic likeness of you for a work meeting or a stylized character for a game.

As you move about the metaverse, your identity, your personas, and your avatars should belong to you and travel with you. This is starkly different from how these components are handled on the web today, where they are re-created and locked inside every platform you enter.

Metaverse browsing lets people move freely among spatial places and services

Metaverse browsers have all the same properties as web browsers, except applied to 3D locations rather than 2D information. On today's web, you open a web browser and move from page to page without ever thinking about who built each site or where it is hosted. Metaverse browsers provide that same freedom within the metaverse. In the real world, you move through physical space, and the services, information, and people anchored nearby appear automatically, with no app to install at each step. Between virtual spaces, you move just as easily, stepping from a meeting room into a game, from a concert into a friend's living room, without logging out of one platform and into another. A web browser is how you navigate the web. A metaverse browser is how you navigate the metaverse.

The metaverse connects these technologies into one open system

The metaverse is the connective layer, the evolution of the World Wide Web, that makes the physical world addressable, shareable, and interactive. It is the open network that ties these pieces together, seen through XR, enabled by spatial computing, acted upon by physical AI, informed by digital twins and world models, and located by positioning. It is to places what the web is to pages. You may also hear the term *spatial web* used to describe the same idea. In this book, that is what we mean by the metaverse.



Chapter 4

AI, Robots, and Smart Glasses Are Converging

> ***"The ChatGPT moment for physical AI is here, when machines begin to understand, reason, and act in the real world."***
>
> **— Jensen Huang**
> *founder and CEO of NVIDIA*

Three breakthroughs are converging at once

Every few decades, something rare happens in technology. Several independent breakthroughs, each created by different people and organizations, solving different problems, converge at the same moment and on the same product. Most people don't see it forming as it comes together, because they are looking at each one in isolation. They might see a robotics story here, an AI story there, and a gadget announcement somewhere else. They file each one in a separate mental folder. Then suddenly, a new technology materializes when one company has the insight to put them all together.

We are witnessing one of those moments right now. But unlike the metaverse hype of the early 2020's, this is much more than a series of news stories about what might happen in the distant future. It is the product of tens of billions of dollars that are being spent at the present time, by some of the biggest companies and investors in the world. Three forces are driving it.

Physical AI is learning to understand three-dimensional reality

The first force is physical AI. An enormous amount of capital and talent is now betting that the next leap in AI is not a better chatbot, but AI that understands and operates within three-dimensional space.

Take World Labs, for instance, the company founded by Fei-Fei Li, one of the most respected figures in AI. World Labs raised $1 billion in February 2026 in a single round. Its backers included the chipmakers NVIDIA, AMD, and the design-software giant Autodesk, which alone put in $200 million.[13] The company's product, Marble, turns text, images, or video into persistent, explorable 3D worlds.[14] Later that year it acquired a robotics-simulation company to close the loop between generated worlds and robots that learn in them.[15] In September 2026, AMD, one of those early backers, agreed to buy the company outright for about $8.2 billion, citing physical AI and robotics.[16] That is not an entertainment bet. It is a bet on the real world coming online.

Li put the goal this way, "If AI is to be truly useful, it must understand worlds, not just words. Worlds are governed by geometry, physics, and dynamics, and reconciling the semantic, spatial, and physical is the next great frontier of AI."

World Labs is not alone. The race is on to give AI a grip on physical space. NVIDIA's Cosmos, a family of open *world foundation models* built for physical AI, has been downloaded millions of times by robotics and self-driving teams.[17] Google DeepMind's Genie generates interactive 3D worlds you can move through in real time.[18] Runway, better known for generating video, released a model aimed squarely at training robots in simulation.[19] And Yann LeCun, one of the handful of researchers often called the godfathers of modern AI, left Meta to start a company on the same conviction, that the next era of AI must understand the physical world, not just describe it.[20] When several

teams of high caliber people all start chasing the same idea at once it's not merely happenstance.

Robots entering human spaces need shared spatial context

The second force is robotics. The investment money being poured into building more capable, autonomous robots further signals something larger than an industry trend. Venture funding for robotics startups hit a record $13.8 billion in 2025, nearly double the $7.8 billion of the prior year.[21] Humanoid robots drew roughly $3.5 billion of that, more than the sector had raised in the previous six years combined.[22] Goldman Sachs projects the humanoid market alone will reach roughly $38 billion by 2035.[23]

No name in the field looms larger than Tesla. It is one of the most valuable companies on earth, valued at more than a trillion dollars, and it has made its Optimus humanoid robot a centerpiece of its entire future. Analysts attribute hundreds of billions of its market value to that single program.[24] Figure AI, building general-purpose humanoids, reached a valuation of $39 billion.[25] Skild AI, which builds a robot *foundation model* (a single brain meant to control many different robots), raised $1.4 billion in January 2026 at a valuation above $14 billion.[26] Physical Intelligence, another robot-brain lab, was valued at $5.6 billion and was reportedly in talks to raise at more than $11 billion.[27]

1X Technologies offered a $20,000 home humanoid called Neo for preorder and reportedly sold out its first year of production within days.[28] Apptronik's Apollo robot entered pilots with Mercedes-Benz and the logistics giant GXO.[29] In China, Unitree shipped more than 5,500 humanoids in a single year, competing on price and volume, before listing publicly at a valuation near $9 billion.[30]

Autonomous vehicles are on the same trajectory. Waymo raised $16 billion in early 2026 at a valuation of $126 billion.[31] Tesla is building its robotaxi on its consumer fleet. Amazon launched a purpose-built driverless vehicle through Zoox. Mercedes-Benz, Nissan, and Stellantis are backing the UK startup Wayve. Goldman Sachs projects the global robotaxi market alone will reach roughly $400 billion by 2035.[32]

These lofty valuations signify that a real platform shift is taking place. Robots that were traditionally fixed in place suddenly find

themselves untethered and navigating the real world. These robots and vehicles all share one problem the moment they leave the lab – they must operate in shared human space. Whether it's a robot on a factory floor, a self-driving vehicle on a city street, or a humanoid in a hospital, each needs to perceive its surroundings, know what objects and services are present, and coordinate with the people and systems around it. That is not a robotics problem to solve company by company. It is a spatial problem that requires global cooperation between companies to solve. The same problem needs to be solved for every pair of AR glasses.

Smart glasses are becoming the primary interface to the spatial world

The third force is a new class of spatial devices. After years of false starts, AR and AI smart glasses have crossed from curiosity into a genuine market, and the world's largest companies are jockeying for position within it.

Meta holds around 73 percent of the AI-glasses market and, together with EssilorLuxottica, shipped roughly 7 million smart glasses in 2025, with plans to reach 10 million units of annual capacity.[33] Its Ray-Ban Display model, at $799, pairs with a wrist-worn neural band that reads the small electrical signals sent from your brain to your hand so you can control the display with subtle gestures.[34] Its newest device for virtual and mixed reality, Meta VR Glasses, weighs about 100 grams, a fifth of its Quest 3 headset, and looks more like a pair of glasses than a headset.[35]

While Meta is the largest player in this space, they are not without competition. Snap's Specs go further. They are full AR glasses at about 130 grams, with a full-color display in the lenses and no phone or tethered puck, so the entire computer sits in the frame.[36] Its CEO called them, "the beginning of a new era in computing."[37] Google and Samsung launched smart glasses on a shared Android XR platform, built with eyewear brands Warby Parker and Gentle Monster so they look like something you would actually wear.[38] Apple is reported to be close behind.[39]

Note the stark contrast between the rollout of these new spatial devices and every other technological product that preceded them. These companies are not competing on sensor counts or resolution. They are competing on whether the glasses look normal enough to wear in public. The lead marketing metric is the brand of eyewear rather than the chips and technology that sits inside them. Adoption, not capability, is potentially the biggest obstacle holding back a product on the verge of going mainstream.

Most of these devices, regardless of the brand, run on chips from the same company. Qualcomm's Snapdragon processors power Meta's Ray-Ban glasses and Quest headsets, Snap's Specs, and the Google and Samsung glasses. By the middle of 2026, Qualcomm counted more than 60 million XR devices in the market. Roughly 600 million pairs of glasses are sold around the world every year, and only about 1 percent of them are smart glasses. Qualcomm's CEO, Cristiano Amon, expects that to change quickly. Within a couple of years, he said, smart glasses could reach hundreds of millions of units a year and "could become as big as smartphones."[40]

These independent markets are rebuilding the same infrastructure

Three forces are converging at once, each from a different direction. Physical AI is learning to understand, generate, and navigate three-dimensional reality. Robots and autonomous vehicles are moving into shared human spaces. Smart glasses are giving people a way to see and interact with a digital layer over the physical world. Beneath each sits an underdeveloped layer of spatial infrastructure that includes billions of IoT sensors, pervasive high-bandwidth networks, and digital twins that turn individual buildings and machines into live, queryable environments. That infrastructure is the metaverse.

The robot and the person in glasses need the same understanding of the same room. The world model that reconstructs a factory and the digital twin that runs it are describing the same physical place. The AI agent that wants to act somewhere needs to discover what services and information are present. Each of the companies that build these products, working in relative isolation and secrecy, are building their

own private version of the same infrastructure that serves the same purpose – to represent the physical world, know precisely where you are in it, anchor information and services to precise locations, and connect to them.

One open metaverse must connect people, robots, vehicles, and agents

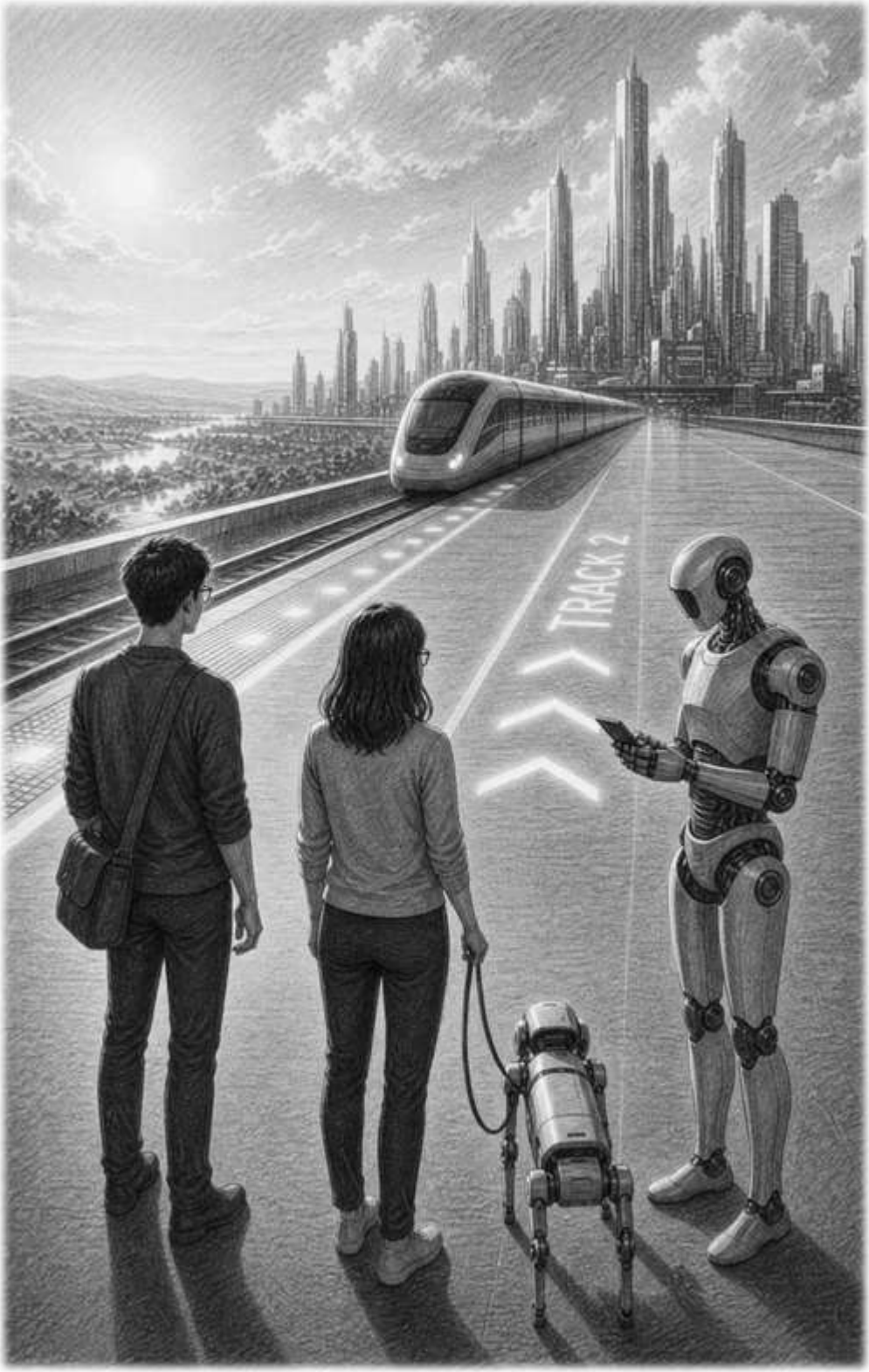
At least one company has recognized this convergence and figured out that the robots, world models, digital twins, and smart glasses are all part of one market. NVIDIA, the company supplying the chips to nearly everyone in this race, is now building across every layer of the metaverse at once: world models in Cosmos, simulation and digital twins in Omniverse, and a framework it calls XR AI that puts perceptive AI agents directly onto AR glasses.[41]

That is not just a loose collection of vertical product strategies. Every layer of the metaverse from the AI running on the glasses, the world models training in the cloud, and the digital twins simulating factories and cities, runs on GPU compute. Billions of always-on spatial devices and the infrastructure behind them represent an enormous amplification of demand across NVIDIA's entire business.

The metaverse is the shared foundation that binds the robots, the glasses, the AI, and the digital twins together. The question is no longer whether or when these technologies will arrive. It is whether they will all connect through a unified open foundation like the web, or fragment into dozens of incompatible private ones.

Part 2

Spatial Computing Is Coming of Age


TRACK 2

Chapter 5

An Open Foundation Could Unlock Smart Glasses

> ***"More than a billion people worldwide wear glasses today, and it seems highly likely that these will become AI glasses over the next five to ten years."***
>
> **— Mark Zuckerberg**
> *co-founder and CEO of Meta*

The phone's form factor separates digital information from the physical world

More than half of everyone alive carries a smartphone.[42] It has swallowed the camera, the map, the wallet, the newspaper, the television, and the telephone that gave it its name. By almost any measure, it is a triumph. It also has three flaws that no amount of success can fix: it makes you look down, it ties up one of your hands, and everything it shows you must fit in a small, flat rectangle.

Because of those flaws, the device that connects us to the rest of the world simultaneously disconnects us from the world that is right in

front of us. To use your phone, you must break contact with the physical world as you drop your gaze and focus on a screen that lives apart from the room you are standing in. We have built an entire civilization of people looking down. It is so commonplace that we have stopped noticing how strange it is, unless someone walks straight into a lamppost or steps off a cliff.

This is not a complaint about screen time – that's a book in an entirely different section of the bookstore. It is a complaint about form, however. Mobile phones force you to choose between the world around you and the information on your screen. You cannot have both at the same time, which is why all but four countries have laws against operating your phone while you drive. That is not a problem that can be fixed with a software upgrade. To fix it, the form factor of the device must change.

Smart glasses remove the phone's fundamental physical constraints

It was never that we loved looking down. It's just that looking up had not been an option – until now. Smart glasses fix all three fundamental flaws of the phone because they overlay information directly in front of you as you move through the world, not on a separate screen. When your display sits in your field of view and your device sees what you see and knows where you are, you stay present within the world around you rather than disconnect from it. The world is your screen.

That is why smart glasses are not a phone accessory the way smartwatches are. They are a candidate to replace the phone as the primary computing device, the first serious one since 2007.

Spatial awareness and co-presence – not floating screens – justify the new form

Smart glasses are not a slam dunk. One mistake could derail the first serious wave of consumer glasses. If the glasses merely take information that is already on your phone screen and float it in front of your face, people will likely conclude that a rectangle hovering in space with highly inefficient input is not a worthwhile replacement for

the device in their pocket. The novelty factor alone will not replace a device that has been perfected over decades. *The phone, but on your face, and with worse input* is not a winning product.

The advantage that justifies glasses, the one a phone cannot match, is that the device intrinsically understands the space around you. This stems from two features. The first is spatial awareness – information is not just shown to you but placed, anchored to the exact machine, doorway, product, or street it belongs to, and it stays put as you move around. The second is co-presence – other people, wherever they physically are, appear with you in your surroundings, seen and heard, instead of squeezed into a grid of video tiles.

Having a colleague stand beside you as you work on a machine you are both looking at, pointing to the exact part that failed while the repair steps sit anchored to it in space – that is something a phone cannot do nearly as well as a pair of smart glasses. Use cases like this are what will cause smart glasses to legitimately replace phones.

Everyday tasks become hands-free and rooted in place

Picture an ordinary day, a few years from now.

Walking to the train station, your route appears on the sidewalk ahead of you, and the arrival time sits over the train platform as you reach it. A distant relative joins you for dinner at the table, sitting in one of the empty chairs. You shop online for a sofa and place it full-size in your living room, trying out various fabrics to see which one matches best before you buy it. A leaky pipe is inspected by a plumber from across town, who can see the pipes behind your wall and direct you to the shutoff valve.

None of those experiences would be natural enough with a phone to make them practical. A whole sidewalk of people, arms out, phones waving in the air, panning around like tourists filming a landmark? Unlikely. Glasses are far superior to phones for spatial interaction. You can still take your glasses off when you don't need them, the same way you pocket a phone. But while you are wearing them, what you need is simply there in front of you.

Smart glasses shipping today have not reached this potential yet. They can take photos, play audio, and respond to voice commands, but

they cannot yet place information in the world around you or let another person appear in the room. The ultimate goal – a bright display, wide field of view, spatial awareness, real-time rendering, all-day battery life – requires fitting the capabilities of a mobile phone into a frame light enough to wear from morning to night. That is today's most challenging consumer electronics problem, and the world's biggest companies are spending billions to solve it because there is a massive windfall in revenue for the company that gets there first.

Enterprise value will fund the technology before consumer adoption takes off

Nearly every major computing platform started with enterprises. The personal computer entered the office before the home. The earliest mobile phones were tools for field executives and salespeople who could justify the cost. The BlackBerry began as a business instrument, adopted because it paid for itself, long before it became a consumer product. Consumers wait for a device to be cheap, small, and pleasant to use. Enterprises do not need to wait nearly that long, because their math is different. If a tool saves a worker time, increases safety, prevents a costly error, or lets one person be in ten places at once, it pays for itself immediately, awkward or not.

Smart glasses will undoubtedly follow the same curve, and for the exact same reason. With each and every pair of smart glasses deployed, training time compresses, errors fall, and downtime shrinks. A headset that looks a little clunky on a factory floor is not a dealbreaker. A five-figure reduction in downtime is what matters more than appearances. This is why the first at-scale deployments of spatial computing are landing in hospitals, warehouses, factories, and airports. Not because those places love gadgets, but because the savings are immediate.

The consumer wave will follow. As glasses get lighter, smaller, and cheaper, pushed along by enterprise demand and by people who have grown accustomed to wearing them at work, they will eventually cross the threshold where a normal person prefers to wear them in place of carrying a phone. Enterprise money funds the R&D that makes the glasses good enough for everyone else. Business first, where the savings are undeniable, then consumer adoption follows soon after.

Carriers have strong incentives to subsidize an always-connected device

The wireless carriers have just as much at stake. A phone gets pulled from a pocket a few times an hour. Glasses worn all day are connected all day, streaming spatial data and handling co-presence sessions that demand far more bandwidth than a voice call or a video ever did. That makes always-on smart glasses the biggest growth opportunity for carriers since the smartphone. 5G already delivers for streaming and everyday mobile use. Always-on spatial computing pushes demand to an entirely different level, and 6G, arriving by the early 2030s, will push it further still.

Carriers will subsidize glasses the way they subsidize phones today, bundling them into plans and financing them at the point of sale, earning back the investment on the data those glasses consume every waking hour. For carriers, the incentives are clear. A new device cycle, more data per user per day, next-generation networks with a massive new workload to justify them, and a front-row seat in distributing the next platform. Verizon, AT&T, and T-Mobile are not bystanders in this transition. They are infrastructure partners with strong financial incentives to make it happen.

"Does radio ring a bell?": Dismissing Today's Glasses Repeats Early Misreadings of the Internet and Mobile Phone

In a 1995 television interview, Bill Gates tried to explain the internet to David Letterman. Letterman recalled that someone said that you could listen to a baseball game using your computer. He quipped, "Does radio ring a bell?" The audience roared – why would anyone need a computer to do what a radio already does? Gates countered that the internet lets you listen whenever you want. Letterman retorted, "Do tape recorders ring a bell?"[43] It was a fair joke at the time. Nobody in that audience could have foreseen what was coming. The internet went on to change how the entire world works.

The mobile phone received a nearly identical reception. The first ones were not pocket devices. They were bolted into cars or lugged around like bricks, and you could only make phone calls with them of course. Plenty of sensible people said no rational person would ever

want to carry a phone like that around. Then the brick shrank, year after year, until it slipped into a pocket and in so doing, it became a device almost no one could imagine living without.

Which brings us to where we are today with smart glasses. The people who cannot picture wearing them all day in place of a phone are potentially making the same mistake at the same stage.

A general-purpose computing platform does not need one killer app

One question comes up repeatedly regarding smart glass adoption. *What is the killer app?* That's the wrong question to be asking and doing so reveals a somewhat shortsighted vision for the future of smart glasses as an accessory that needs a reason to exist, rather than a platform that will transform everything. Asking for the killer app of smart glasses is like asking, in 1993, for the killer app of the personal computer, or in 2007, for the killer app of the smartphone. There never was just one. Both the computer and the phone gained ubiquity because they were devices on which every app functions.

Smart glasses are not a product in search of a feature. They are the next general-purpose platform and successor to the mobile phone. They will do what the phone does, and many things it cannot. Every service you interact with today will run on smart glasses, presented directly in your field of view instead of boxed in a rectangle. It is how people and organizations will one day interact with the real world.

A pair of smart glasses that can reach the web inherits everything already on it. The use cases are too numerous to describe, which is why no single one use case could ever be the answer. So when someone asks what smart glasses are good for, like the cell phone, the answer is the same – nearly everything. And that is exactly how the phone gets replaced.

Glasses can fulfill their promise only by browsing an open spatial world

This is all a moot point if smart glasses have nothing open to connect to. A pair of smart glasses, by itself, is a camera, a speaker, and a small screen in your field of view. To fill that field of view with useful

information, they need to be able to discover information and services anchored to the physical places you move through. They need a map of the real world to browse to be truly useful. They need the metaverse.

As an evolution of the web, the metaverse is by definition open. Any pair of smart glasses equipped with a metaverse browser can reach anything that anyone else chooses to publish to a location, and without permission from a gatekeeper. When that comes to fruition, the ordinary day above becomes possible.

If, on the other hand, each smart-glasses manufacturer, instead, were to build its own closed world, reachable only by its own devices, then the future fractures. Your directions may work but your colleague's tools may not. One brand's device may see a restaurant's menu but another brand's device may not. The vision for the ordinary day falls apart as each device is locked out of dozens of incompatible islands.

With spatial devices, robots, and physical AI well on their way, the trillion-dollar question we must answer now is, *what will they be allowed to connect to?* Dozens of proprietary, walled-garden ecosystems – or a single open metaverse born from the web that has been the rock-solid foundation and the greatest source of innovation over the past four decades. The metaverse is the only reasonable answer, and undoubtedly the one, like the open web, that will be impossible to compete with.



Chapter 6

Co-Presence Could Replace Calls and Trips

"Using this instrument, you can 'work' in another room, in another city, in another country, or on another planet."

— Marvin Minsky

co-founder of the MIT Artificial Intelligence Laboratory

Co-presence preserves the value of a business trip without the travel

Business trips exist because someone from one company needs to be physically present with a client or a partner in a boardroom or on a job site. Most trips involve a flight, a hotel, a few days of travel, and an employee diverted from everything else they could be doing. The cost is not only counted in dollars spent, it is also counted in time, fatigue, and lost opportunity. Then there are the missed connections that never happen at all because the trip was too costly to justify.

Companies spent $1.59 trillion on business travel in 2025 and are on track to spend $1.71 trillion in 2026, with the global total projected to pass $2 trillion by 2030.[44] That is the budget line co-presence competes with, and precisely why enterprises are eager to adopt this technology long before consumers.

Corporate budgets are not the only thing under scrutiny. Air travel accounts for around 2.5 percent of the world's carbon emissions, and many large companies have now signed verified targets to cut business travel, some of them in half.[45] That leaves companies with two options. Send fewer people and lose out on the benefits of the trip, or find another way to get the net effect of an in-person meeting without the travel.

Getting that effect is more nuanced than it may seem, because a trip is not exclusively about the conversation. If talking were enough, a simple phone call would suffice. People make trips for two main reasons. One is to strengthen a relationship, which comes from sharing surroundings and an experience. The other is to be on premises, because some things must be seen in person. A video call serves neither.

Co-presence, on the other hand, allows you to virtually join someone at a real location so that you don't have to physically travel there. All it requires is a phone, computer, or preferably, a VR headset. The people present at that location can see and hear you, and you can move around and talk with them in much the same way that you would if you had been there in person.

While it may never be possible to replicate an in-person meeting with one hundred percent accuracy, co-presence promises the next closest thing to reality and provides many more benefits than a video call. Two or more people share a lifelike experience in the same surroundings, which is how relationships are built. You can stand beside someone while you talk and gesture, turn together to look at what one of you noticed, and stay around to socialize after your business is done. All while preserving your company's bottom line and your valuable time.

Digital twins and avatars let remote people share a real place

The underlying foundation of co-presence is a digital twin of the real location in the metaverse. People who are present at the physical location wear smart glasses and exist inside the digital twin as they move through it. Those who are remote can be virtually present in the same location from wherever they are via a VR headset, a desktop computer, or even a phone. Their device renders the full 3D representation of the location along with anyone else present in the space, either in person or remote.

All participants can see, hear, and point at the same objects, provided they are represented in the digital twin. The experience for the remote attendee is strongest in VR, because it allows them to easily move around in space and see things naturally by moving their head, much the same way they would if they were actually present in the location.

All attendees are tracked with six degrees of freedom, which in practice means head, face, hands, and body are all updated continuously. Gaze, posture, pointing, and position within the location all contribute to a very realistic and rich communication experience.

Those present in the real-world location will continue to see others who are also present at that location the way they always have – looking through the lenses of their smart glasses as one would look through a normal pair of glasses. For everyone else, people are represented by their avatar. This is true for both the in-person smart-glass attendee looking at someone attending in VR, as well as the VR attendee looking at everyone else.

Spatial audio contributes more to presence than photorealism

Avatars can take on a variety of appearances, from heavily stylized to incredibly lifelike. In fact, the technology has advanced so much in the last several years that you can have a real-time updating avatar of yourself that is so close to your actual likeness, hardly anyone can tell the difference from across a table. It turns out, however, that your avatar appearance matters less than you might expect.

Research shows that sound matters just as much as appearance when it comes to co-presence. In a controlled comparison, adding

room acoustics to head-tracked audio did as much for the sense of being present as multiplying the visual resolution by five.[46] A photorealistic room with flat stereo audio feels like watching a screen. Whereas, a plain room where voices arrive from the correct direction, footsteps land behind you, and the hum of a machine three aisles over shifts as you turn your head, feels more immersive.

Spatial audio is also what makes a group conversation seem realistic. When six people are in a room and each voice comes from where that person is standing, your brain intrinsically knows who is speaking. Conversations happen more naturally. You turn toward people you are listening to, and two people can talk quietly on the side without derailing the main conversation. Flatten those six voices into the center of everyone's head and your brain will perceive it as it does a conference call. This works just as well in a one-on-one conversation, a full shift crew, a lecture hall, or several thousand people in a stadium.

Existing immersive meeting products failed on place and friction

Anyone who has been around this industry for a while likely has a healthy dose of skepticism when it comes to the future vision for smart glasses. Meetings in headsets already exist. Meta built Horizon Workrooms, Microsoft put immersive spaces into Teams, and Apple shipped Spatial Personas on the Vision Pro. Billions of dollars have been invested into developing and marketing them, yet none of them seem to have made a material impact.

Each of those products stalled for a similar set of reasons. Foremost, none of them could provide the type of co-presence that puts you into a real-world location where work is taking place, because doing that requires a live model of a real location reachable by any device. Instead, they each dropped attendees into a generic virtual meeting space. That experience competes directly with video calls, which have considerably less friction.

Friction always plays an outsized role in adopting new technologies. A zoom call is one hyperlink away, which is a very low-friction entry point that anything replacing it must be able to replicate. Zoom calls, like phone calls, are available on any device through a

browser and without account setup, so there's never any question of who has the right hardware or software setup to participate. Because each of these three products exists only within a closed, walled-garden ecosystem, only those who have the right gear, a preinstalled app, and an account can participate. Every additional step loses people. The more steps between a person and the place they need to be, the fewer people arrive.

It is vital that we don't walk away from this experiment drawing the wrong conclusion. What failed was not the notion that co-presence, done correctly, can be an effective replacement for a business trip. What failed was a generic meeting room hosted on a closed hardware ecosystem, accessible through significantly more steps than a video conference call that is accessible to everyone.

Co-presence reduces video fatigue and expands remote work

The Covid pandemic of 2020 settled the question of remote work by accident. Companies that had never allowed remote work were given no choice, and most of them found they not only could still operate, but they could do so more efficiently. Business deals closed over video calls. Products were delivered. Travel budgets went unspent. The net result – a Stanford study published in Nature found no negative effect on productivity from hybrid work, along with a drop of a third in employee turnover.[47] The experiment ran at global scale, and the remote work question found its answer.

Some companies wasted no time leaning into the remote work model. In early 2020, for instance, Pinterest paid $89.5 million to walk away from a 490,000 square foot San Francisco office it had leased before the building was even built, citing the shift to remote work. Cancelling their lease cost the company far less than the roughly $440 million in rent it would otherwise have owed.[48]

At the time, the video call was the best tool available for conducting work remotely. Six years removed from the pandemic, it's still the most common tool in use today for meetings, even though the considerable strain it puts on people has been documented. Researchers at Stanford who studied why videoconferencing is so exhausting found the largest predictor was feeling physically trapped,

held inside the narrow cone of the camera lens. The second was looking at yourself all day, which no prior medium required people to do.[49] Anyone who has participated in a video call with several faces holding direct eye contact a foot from your own, regardless of who is speaking, has undoubtedly experienced this fatigue.

Co-presence removes those problems instead of easing them. You are free to stand up and walk around, because your body is the input. You never see your own face, and attention is given to whoever is talking, the way it does in a normal conversation. While these factors can't fully eliminate the office, they mitigate the main reasons that companies give for requiring one.

Being in the same room is only the start of things that co-presence can handle. Digital twin offices can also provide services that substitute virtual tools for most of the things people rely on at work such as whiteboards that anyone can write on, full scale product models, and shared screens – as many as you need. These services don't even need to be installed. They are provided as part of the workspace the same way the furniture is, and anyone who walks in can use them.

Consumers gain presence for trips they could never take

A company can determine if a business trip is worth the cost and send someone if the return on investment is positive. Families, on the other hand, operate on a different set of measures. Out of town grandparents, for instance, might only get to see their grandchildren twice a year because flights are too expensive. Others may not even have a choice, such as a parent serving in the military stationed overseas for nine months. In such cases, people will settle for an occasional phone or video call.

Roughly one person in ten flies in a given year, which means replacing trips only serves as a motivating factor for a small slice of the world's population.[50] Therefore, the primary use of co-presence for the average consumer is to include people who never had the time or the means to be present at all.

Nothing in the technology changes between the boardroom and the family room. Enterprises will pay first, because they have the budget and the savings are easy to measure. Consumers represent a

larger user base, but their reason for using the technology has less to do with cost and more to do with staying connected. Being far from the people you love carries a price no budget records, and no technology to date has done much to lower it.

Smart glasses need interoperable co-presence

The iPhone accounts for half of Apple's revenue, more than $200 billion a year.[51] Apple, Samsung, Google, and a dozen other companies have built their businesses on the assumption that the phone is the device people reach for. If smart glasses eventually become the way people connect with each other, the phone will stop being the primary device. Every one of those companies knows it, which is why they are all racing to build glasses. It is not a side project. It is survival. The first company that builds a pair of smart glasses that can replace a phone will enjoy a huge market advantage.

If those smart glasses were only available inside walled gardens, co-presence would fracture along corporate lines. Your glasses, for instance, may not be able to join your daughter's headset because they came from different makers. Co-presence would come with an asterisk – together, so long as you both bought devices made by the same manufacturer.

The telephone industry settled this conundrum a century ago. A phone call works regardless of your carrier or your handset, because the industry built it on open standards that let any device reach any other. The telcos did not lose by opening up. They thrived. The smart glasses industry has not learned this lesson yet. They will soon enough. Once the rest of the world standardizes on co-presence that runs on an open metaverse, it's unlikely that any closed ecosystem will be able to survive.


375 F

Chapter 7

Spatial Services Stream In Around You

> ***"Imagine you walk into a room, and the room is dynamic… you are interacting with the things going on in the room."***
>
> **— Eric Schmidt**
> *former CEO of Google*

The physical world breaks the app model

Ever since Apple opened the App Store in 2008, and with the Google Play Store arriving right behind it, we have organized our digital lives around single-serving websites, otherwise known as apps. There are roughly five million apps across the two stores today.[52] For those that you use repeatedly on your mobile phone, you search for and find the app, install it, and then open it. The average person uses about ten a day and thirty a month.[53] This process works well enough because our digital life is managed on a small screen, one app at a time, and on our own schedule. The app model has become so familiar that it almost

feels like the natural order of software rather than the consequence of one or two companies' quest to generate revenue. However, that model is about to meet something it was never meant to handle – the physical world.

The app model fails as people move through short-lived physical contexts

The app model rests on an assumption that the physical world breaks. It assumes you go to the software. You seek it out, set it up once, and return to it again and again. The physical world relies on the opposite assumption. It requires that the right software be automatically available the moment you arrive at a particular location and automatically removed when you leave. You are not going to set up an account with a vending machine in a lobby you only visit once. You just need it to work while you stand in front of it.

As you move through a real, spatially aware environment wearing smart glasses, the things that matter to you change every few steps. A storefront here, a machine there, a gate, a room, a vehicle, a kiosk, each with its own relevant information and functionality. Try to imagine how that would work in the metaverse using today's app store model – stop, figure out which app you need, find it, install it, open it, grant it permissions, agree to its terms of service, and create an account. Then do it again ten feet later. And again, and again, and again. Nobody wants to do that.

As a thought experiment, let's consider some real locations. Walk through a shopping mall and count all the apps you'll encounter along the way: fifty stores, twenty restaurants, a parking garage, a movie theater, a few kiosks, and a pay station on the way out. Under the app model, that is dozens of downloads in a single visit. Now run the same thought experiment in your mind for an airport, a hospital, an office park, or a university campus. The friction we tolerate a few times a month on a phone becomes absurd when the world around you is dense with information and services that you might want to interact with.

The app model is not merely annoying for smart glasses, it would undermine adoption because the primary drivers of adoption factors

are friction (or lack thereof), interoperability, and whether the experience is actually spatial. If interacting with the things around you requires a download and install every few steps, or if a service only works on one brand of device, people won't use the glasses. Similarly, if they merely provide a 3D display for 2D information, people will not see the point – their phone already does that, and it does it better. Those expensive smart glasses will end up in your drawer next to your discharged VR headset and the industry that spent billions on building and marketing these devices will wrongly conclude the world was not ready for them. The world is plenty ready. The experience will simply have been too cumbersome.

The lack of friction is exactly why the web browser took off. Jumping from one website to the next carries no friction at all. You click a link and a page loads with no install, no account, and no regard for whose site it is or what device you are using. That near-zero friction, not the novelty of the technology, is what turned the web into something everyone used. The same rule governs the metaverse. If we want the world to adopt spatial computing, interacting with a service must be that easy, and it must work everywhere, and on every device.

Spatial services replace installed apps with proximity-based, sandboxed software

When moving from mobile phones to smart glasses, the model must invert. Instead of apps you install and run on a whim, the physical world runs on services that surround your current location. A spatial service is a piece of software anchored to a place or an object. Your device discovers it by proximity, connects to it instantly, uses it, and unloads it as you move away. Like loading a webpage, there is no manual download or install process. But unlike a webpage, you can have many services from different providers running concurrently, composed into one scene in front of you. The shift is from a world of apps you manage to a world of services that find you.

Services come in two basic categories. Most services fall into the first category in which they are anchored to an object or location. Examples of this might include a wayfinder at an airport or a diagnostics panel on a printing press. The service is placed there by the

owner of the object or location, and it activates automatically whenever you get close enough to interact with it.

The second category of service is built to travel with you. These are the tools you want to use wherever you go, and they belong to you rather than an object or location. Examples of this type of service include a translator, digital wallet, or games. One such service that is bound to be exceedingly popular is an AI assistant, which will be fully capable of seeing and interacting with the metaverse around you. It is way more than a chatbot on a screen. It can sense the space around you, inform you, guide you, and act on your behalf if you choose to allow it.

All services in both categories run sandboxed within a metaverse browser. This follows from the core property of web browsers that also sandbox JavaScript code so that it cannot harm your device. This is vital because software that arrives unannounced and retrieved from a server your device has never encountered is entirely untrusted. It must be contained so that it can see and do only what it has explicitly been granted.

Earlier walk-up technologies failed because they delivered links, not services

Walk-up software such as QR codes, NFC tags, Apple's iBeacon, and Google's Eddystone already exists. Google went the furthest, building proximity notifications into Android itself in 2015 so that a phone could pick up information from a nearby beacon with no app installed at all. Three years later Google switched it off because the notifications had turned into spam.[54]

Those systems delivered a link, not a service. A beacon knew you were near it and nothing else, so the best it could do was push an alert and hand you off to a webpage or an app install. Two beacons in the same store were two interruptions, not one composed scene.

QR codes survived because they are transparent about their intent and require you to initiate the interaction. A QR code is still a static link, though. It knows nothing about the space it is affixed to, and it will never alter its behavior based on what is happening around you.

Services, on the other hand, actively provide real-time utility related to the object or space to which they are anchored. There are two forces that regulate how prominent they are when you approach. The first one is your own browser settings, which allow you to customize what happens when new services activate. You can filter services by properties such as type or publisher. And your own personal AI agent will learn which types of services you do or do not wish to interact with.

The second force at play relies on businesses and municipalities to act in their own best interest, just as they do today. Services deployed by a business are designed to serve their customers in the same way they arrange their customer's experience now – just virtually. If you find them obtrusive, it will be a simple matter to turn them off for that location. When too many customers do that, the business will most likely respond by improving the experience. This is how free markets tend to work.

Location owners control what appears while providers host their own software

While each company owns and controls its corresponding locations in the metaverse, the same way it controls its own website, the services it offers are often hosted by third-party providers rather than on a company's own servers.

For instance, a coffee shop could drop a virtual backgammon board game onto one of its coffee tables for customers to play while they sip their venti, double-shot, half-caf, bone-dry macchiato. The backgammon server, however, is run by a third-party game company. This enables users to communicate directly with the game vendor, who need not hand over its code or data to a closed platform to operate. This also spares the coffee shop the trouble of figuring out how to build and maintain their own backgammon service, which is outside their core competency.

This same layered integration approach works for larger enterprises like hospitals who can contract with healthcare service provider organizations that develop and operate complex services for thousands of hospitals worldwide. Airports are another example that

perfectly demonstrate how third-party services can be layered and integrated into a complex ecosystem that is tailor made for the metaverse. Take, for instance, a baggage conveyor service that runs on the manufacturer's own servers and is reached directly by the device of the baggage handlers who are within proximity of the machine. This service might be hosted in the cloud on a server that manages a thousand other conveyors in airports across the world. The conveyor company develops and maintains the conveyor service, relieving the airport from having to do so, which is outside its core competency.

The airport has its own reasons to want this across a wide spectrum of services that operate throughout its terminals and on its tarmacs. Flights leave on time when a conveyor belt gets fixed in twenty minutes instead of two hours. Passengers who can more easily navigate security and find their gate miss fewer flights, and the time they save gets spent in shops and restaurants.

Relatively few of the people who use these services are actually employed by the airport itself. More than sixty-three thousand people work at the Hartsfield-Jackson Atlanta International Airport, the overwhelming majority of them for airlines, contractors, and concessionaires rather than for the airport authority, and 106 million passengers came through in 2025.[55] Airline crews, baggage handlers, the conveyor manufacturer's technicians, and passengers all access airport services on devices the airport does not own and cannot manage. This would be all but impossible to facilitate any way other than on open metaverse infrastructure.

Services can extend to every product without app-store friction

Samsung, GE, Bosch, and every other appliance manufacturer that puts a product into the physical world can deploy services in the same way as the conveyor company in the example above. This benefits consumers and manufacturers alike. For the consumer, a washer will show you its controls, its cycle options, and its maintenance history the moment you look at it. Less commonly used features documented in an owner's manual you lost years ago will guide you through the steps, right on the appliance panel, the moment you need it. If you're unsure of how to do something, you can just as easily ask the service a question

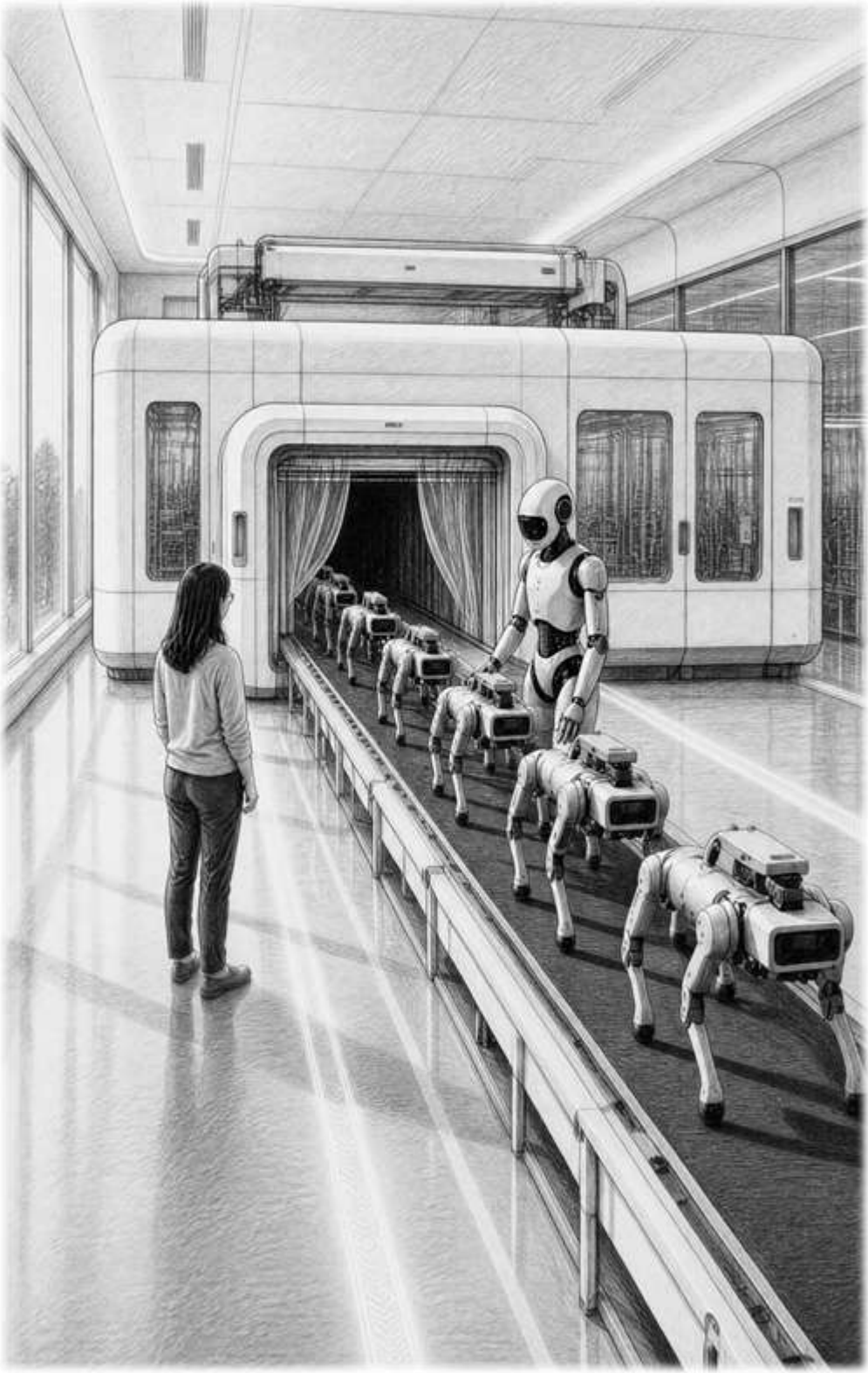
or ask it to change a setting for you by speaking to it in your own language.

For the manufacturer, this is a direct relationship with every customer. Reaching that customer through an app today means building for two platforms – iOS and Android – submitting to each of them for approval, not to mention forking over the thirty percent of all revenue sold through the store, graciously reduced to fifteen percent for developers who transact under a million dollars a year.[56] Manufacturers must also perform continuous maintenance for each app for as long as it exists, because both platforms change policies on occasion. Apps that fall behind or stop working will get removed.

Despite all the things an app developer must do, ultimately, it's still left up to the appliance owner to install and use the app. Only about a quarter of people who install an app open it again the next day, and roughly one in twenty will continue to use it a month later.[57]

A metaverse service has no toll and needs no approval. And there is nothing for the customer to keep up to date. Just like websites, services receive updates whenever the manufacturer updates them on their server. A product that comes with an attached metaverse service is more useful than one that does not. Every business built a website twenty years ago for the same reason. If your competitor's product talks to the customer and yours does not, yours loses.

The result will undoubtedly be hundreds of millions of services deployed in the metaverse by businesses of all shapes and sizes. Contrast that with today's app market. Among the 400 million small and medium-sized businesses worldwide, which account for 90% of all businesses, only about five million apps exist.[58] This shows that the app model mostly serves the companies large enough to afford the development and marketing cost of building an app. Everyone else is priced out, which is why the corner store, the dentist, and the municipal parking garage don't have apps. Yet all three have a website because the web is open and costs comparatively little to participate.



Chapter 8

Physical AI Delivers Spatial Intelligence

> ***"Spatial intelligence – AI's ability to perceive, reason about, and interact with the 3D world – is the scaffolding upon which human cognition is built… It's the difference between describing a room and navigating it."***
>
> **— Fei-Fei Li**
> *co-founder of World Labs*

AI cannot act in places it cannot understand

Today's AI can pass a medical licensing exam and write working software. Ask it how far you're standing from the nearest fire exit, and it won't be able to give you an answer. The question is not a hard one. Except nothing has told the system anything about the room and where you are standing inside of it.

AI learned to read and write first because text and images are abundantly available on the web and easily transmissible, thanks in no small part to all the web's standards and protocols. The physical world

is none of those things. It is continuous, it changes constantly, and it is mostly not online.

An AI software model that integrates with sensors and actuators to perceive and act within the real world is called Physical AI. It takes in what is around it and acts on it, instead of only working with text and images. The term often gets used interchangeably with the term *robot*, but that is too narrow. For instance, an AI assistant running on a pair of smart glasses is physical AI. The term also describes systems such as software that schedules maintenance across a company's buildings, robots that walk a plant at night checking gauges, and vans that drive themselves. Some of these are robots. Most physical AI is not.

About 2.25 million units of smart glasses with a built-in assistant shipped in the first three months of 2026, roughly what the entire category sold in all of 2024. The full 2026 year is on track to sell around 13.6 million. Gartner expects 40 percent of business software to have task-specific AI agents working inside it by the end of 2026, up from under 5 percent a year earlier.[59] Robots are further behind. Fewer than 20,000 humanoid robots shipped worldwide in 2025, and most went to research, education, entertainment, and data collection rather than production work.[60] The factory deployments that do exist put tens of robots, rarely a hundred, on a single task such as moving parts or totes.[61] Four-legged inspection robots are further along, with more than 1,700 from the two best-known makers in customer hands,[62] walking rounds at plants, substations, and offshore rigs.[63]

AI leaves the text box when places can describe themselves

Today's AI assistants possess immense knowledge about the world in general. But they have no knowledge about the room you are standing in unless you make up the difference by describing your surroundings with text or photos. That process disappears once the location you're in can describe itself. A company can do that for one of its office buildings the same way it publishes a website. It puts a live digital twin of its building at an address, on its own servers, for any metaverse browser to read it.

The digital twin identifies each object within its perimeter, including its location to within millimeters, what it is, what it is for, along with who and what people are allowed to do with it. It is not a three-dimensional copy of the company's webpages. It is a live three-dimensional model of the building, and the company decides what goes into it, as well as which parts of it are public, and which are private.

An AI assistant connected to the digital twin can now tell you how far you are standing from the nearest fire exit without you having to describe anything. It can also do things, not just answer questions. An assistant that knows where it is can do location specific tasks such as adjusting a valve on a machine you're monitoring, book the room you are standing outside of, or hold the elevator you are walking toward. All this will be done from a metaverse browser without having to open separate apps and do it yourself.

Seeing objects does not reveal their spatial relationships or state

A good AI model with a camera can identify almost anything in view. Knowing what things are is not the same as knowing where they are in relation to each other. The latter is what these systems still have difficulty with.

Researchers at Stanford and New York University tested exactly this problem. They filmed ordinary indoor spaces, then asked leading AI models questions that anyone standing in those rooms could answer. How far was the sofa from the television? Which of two objects was bigger? What had they passed on the way in? People got about eight in ten right responses correct. The best AI model managed a little better than half of that.[64]

When the researchers sorted the wrong answers by cause, roughly seven in ten had nothing to do with the model failing to see something. It saw the objects and named them correctly. It simply got their sizes, distances, and directions wrong. Asking the model to reason step by step, which reliably improves its answers on written problems, ironically made its spatial answers worse.

Better models are not guaranteed to fix this issue. A system can have perfect vision and still not know that the machine in front of it has been switched off for service, or that a pallet in the aisle ahead will

be gone in ten minutes. That information may be known in a digital twin, but it is not visible in the picture, so disconnected AI models won't have access to those details.

An equally difficult problem is determining an object's size. A photograph of a dollhouse and a photograph of a real house can look very similar. Film studios have used that fact for a century to shoot model ships and toy cities and pass them off as real. This ambiguity occurs with any single camera system. A room and a perfect half-scale copy of that room produce the same picture, so its true size cannot be worked out from images alone.[65]

Accurate measurements require at least one piece of information external to the picture. That can be an object of known size in the frame, a second camera a known distance from the first, a motion sensor, a satellite fix, or a surveyed marker on a wall. Without something of that kind, an AI system can build a perfectly consistent model of a space and have no idea whether it is looking at a warehouse or a shoebox.

This matters when two systems must convey information between each other. A technician's glasses, for instance, might show a shutoff valve thirty centimeters to the left of a gauge. A robot working on the same machine must be able to similarly measure thirty centimeters from the same starting point.

Today's successful robots and self-driving cars depend on environments prepared in advance

The robots and vehicles that are operating reliably in physical space today had their world prepared for them in advance of their deployment. Amazon runs more than a million robots, and every floor they drive on was intentionally engineered around them, down to the shelving.[66] Similarly, Waymo's vehicles carry passengers in cities that were surveyed and mapped in detail long before the service in those cities opened.[67] That preparation is why those deployments succeeded. It is also why there are so few of them.

The same problem with public roads is well documented. Self-driving cars rely on high-definition maps that record lane markings, curbs, and signs to within a few centimeters. Building a detailed map

like that runs around a thousand dollars per linear kilometer of road, but roads change faster than survey fleets can re-drive them. A map that has fallen out of date is arguably worse than no map at all, because vehicles trust it.[68]

The cost of producing a map is not the main problem. Every manufacturer keeps its own maps in its own format. A city that wants vehicles from two suppliers either commits to one of them or pays to have the same streets mapped twice, and then keeps both copies current forever.[69] The same problem is bound to happen inside every building that installs equipment from more than one vendor.

Hand-surveying every building on earth is not going to come cheaply, and doing it multiple times for each new supplier adds friction and additional cost. Without something better, physical AI adoption will be severely hampered. Robots will work in the buildings whose owners could afford to survey them in advance, while other buildings will remain off limits.

Robots need precise, verifiable information that people can infer

People don't have a problem when they walk into an unfamiliar building. They read the signs, follow arrows, or ask others for directions. Robots, on the other hand, rely on a constant stream of information that they collect from onboard sensors, or from data that is sent to them, which must be carefully verified.

The 2025 revision of ISO 10218, the main international safety standard for industrial robots, added requirements covering the security of safety-related communication. Anything a robot relies on to keep people safe now must be protected against being corrupted, delayed, lost, reordered, or inserted. Safety settings must be checksummed, and only authorized people can change them.[70]

Those rules were written for data moving inside a single robot cell, built and signed off by one integrator who takes responsibility for the whole thing. Whereas information published for an entire building travels across suppliers and companies that have never worked together, and no single integrator is in a position to verify it. A robot must be able to establish that the safety boundary it just picked up genuinely came from the company that runs the floor.

A robot given the wrong safety boundary can move into a space where somebody is working. That is a strong argument for building this layer in the open. A private integration is only as trustworthy as its supplier decides to make it, and the customer has no way to audit the work. A layer built on published standards can be inspected by the facility, the robot maker, the insurer, and the regulator, and when something does go wrong there is a record of what the robot was told and who gave it the information.

A shared spatial foundation eliminates repeated integration costs

In the absence of open metaverse infrastructure, every robot supplier must work out its own way of understanding each location it enters. That entails a proprietary integration project per building, which only serves that supplier's robots. Connecting robots to software a warehouse already runs on commonly takes three to six months of engineering before the machines do any useful work.[71] That cost repeats for every new site, and again for every new supplier brought into it.

Any company buying robots carries its own version of that cost. Their investment is tied to whichever supplier did the integration, and bringing in a second supplier means potentially duplicating that cost.

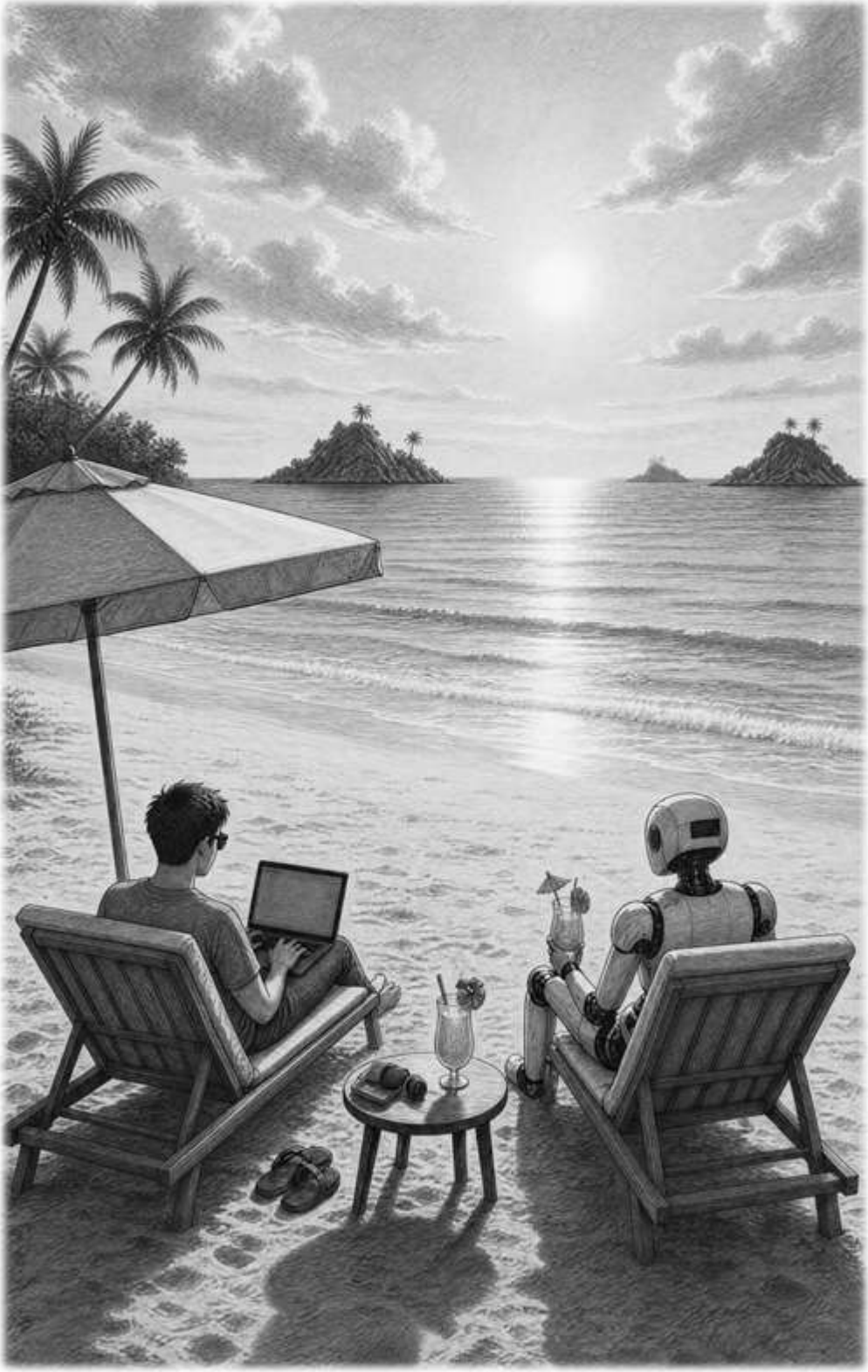
A facility that publishes its own environment once removes both of those proprietary costs. Humanoids on the assembly line, inspection robots on rounds, and wheeled units moving material can all operate on top of the same standardized metaverse infrastructure. The facility keeps ownership of its environment and is free to add or change suppliers. The suppliers compete on how good their robots are rather than on the cost of integration, where the first supplier holds a significant advantage.

It also changes the size of the market for whoever builds the robots. Today a robot can only be sold where its maker has done the integration work, which makes every sale a project of its own. A robot that reads whatever the building publishes can be sold into every building that publishes.

This is exactly what made the web grow so quickly – publishing and reading were separated from each other. A browser does not care who writes the pages it loads. It only cares that they are published in a recognized web standard. Similarly, robots built to read from metaverse infrastructure have no reason to care who constructed the buildings they walk into. They should only care that they have a digital twin published in a recognized metaverse standard.

Part 3

The Web and Walled Gardens Both Fall Short



Chapter 9

It's a Web Browser Problem, Not a WebXR Problem

"The Web as I envisaged it, we have not seen it yet. The future is still so much bigger than the past."

— Tim Berners-Lee
inventor of the World Wide Web

What WebXR is, and what it was built for

Among those who agree that the metaverse is an evolution of the web, many believe the most natural progression would be to simply evolve the web browser. After all, web browsers already have a built-in standard called WebXR that makes it possible to add immersive, three-dimensional scenes to existing webpages. Unfortunately, that instinct lies somewhere between entirely wrong and completely impossible, which has led many capable people to spend far too long working on solutions to the wrong problem.

WebXR is a recent addition to web browsers that lets a single webpage display three-dimensional, virtual-reality, and augmented-

reality content inside an ordinary web browser. It was designed so that one company could deliver one finished experience from one source.

The current implementation of WebXR is not perfect. For instance, it's not supported on all browsers, and where it is supported, it does not always behave the same. Safari does not implement WebXR on the iPhone, the iPad, or the Mac. So, on the most widely used phone in the world, a webpage cannot open an immersive experience at all. Even on the Vision Pro, Apple's own spatial computer, Safari will open a virtual-reality session but not an augmented-reality one. Programs installed on that device can place content around your room, but webpages cannot.

Where WebXR is supported, its capabilities shift from one platform to the next. The Quest browser supports the richest feature set, but Chrome on Android supports a slightly different set. Engine vendors publish compatibility tables to advise developers to test for each capability separately and write a fallback for each one. Camera access, depth sensing, scene understanding, and visual positioning are all fragmented across platforms. What works on Quest does not work the same way on Vision Pro, and neither matches what ARCore or ARKit provides natively.[72] The feature gaps are worse in the capabilities that augmented reality depends on most, which include accessing the camera and reading the shape of a room.

Performance with WebXR is another issue that leaves much to be desired. A headset must maintain seventy-two, ninety, or a hundred and twenty frames a second, depending on the device. WebXR's timing guarantees are softer than those of native applications. Browsers implement runtime security barriers between webpages and the hardware, which introduces a performance cost that native programs never pay. Modern JavaScript engines compile code just in time by interpreting or quickly compiling it, then optimize frequently run code into fast machine code, which adds latency that native XR runtimes avoid. All these plus memory constraints, restrictive threading, and imprecise control over the main message loop contribute to hampered performance, which matters for anything demanding real-time precision.[73]

WebXR draws the picture and leaves everything else to the builder

None of these aforementioned limitations are the reason why the metaverse cannot be architected on top of WebXR. Here's why.

When someone builds a conventional website, the web browser does a lot of the heavy lifting. It handles rendering, page layout, security, navigation, session management, and input among other things. The website builder's primary job is to supply the content. That symbiotic arrangement is probably the web's greatest strength.

WebXR does not extend that arrangement into three dimensions. At its core, WebXR only provides some user input and low-level 3D rendering into a bitmap canvas. For immersive 3D content, it manages two bitmap canvases – one for each eye. From the browser's perspective, WebXR's job begins and ends with rendering pixels. Real-world positioning, maps of the surrounding space, networking for co-presence, identity, avatars, 3D geometry and graphics libraries, ray tracing, shaders, spatial audio, menus, dialog boxes, input controls, virtual keyboards, camera access, VPS, and a plethora of other features that the metaverse requires are all left for each individual builder to solve from scratch.

That wasn't even the bad part.

Three.js, A-Frame, and Babylon.js help with parts of this, but all are third-party libraries that each developer must select and wire up themselves. Because they aren't integrated into the browser, each company ends up building its own platform on top of WebXR just to ship one experience, at close to the cost of a native application, but with fewer capabilities.

To be fair, WebXR programs do have one significant advantage over native apps in that they don't need a separate download and installation. The friction for accessing a WebXR program is the same as accessing a webpage in a browser – next to none.

The web's location information cannot place content

The Geolocation API has been built into web browsers since 2009. It returns a latitude and longitude with an accuracy that ranges from 5 to 100 meters depending on your device. Altitude can come

back empty, so a webpage may not know which floor you are on. There is a value that provides the direction your device is travelling, not the direction it is pointing, and that value is empty whenever the device is not moving. A page can therefore tell which way you are walking but not which way you are looking, and while you stand still it knows neither. Standing still and looking at something is most of what a person wearing glasses actually does in the metaverse.

Your position does not serve as an anchor. For one or more services to attach content to an object, the browser must know exactly where that object is relative to your position as you move about your space. WebXR can track a headset accurately while an experience is running but forgets it the moment the experience ends. Neither the web's location interface nor WebXR gives independent publishers a point in the real world, accurate to a few centimeters, that they can attach content to.[74]

While web browsers have URLs, hyperlinks, search, and a way to ask for your coordinates, neither web browsers nor WebXR provide a way to discover content and services around you, load them, place them in the scene, and drop them when you move out of range. On the web you load a single destination URL and remain on that page until you decide to leave by following a link or performing a new search in your tab. Nothing in WebXR describes content that arrives and departs when you change your location by twenty feet.

We still haven't gotten to the bad part yet.

Every website makes you rebuild your identity

Every website owner today must build its own login/persona/avatar system, so somebody moving from one WebXR experience on one site must repeat the process on the next site. Even signing in can be very awkward because the browser's own trust prompts kick users out of immersive mode.

Your identity isn't the only thing that gets left behind when you move to the next site. Each new site takes you in and out of immersive mode and must ask separately for access to your microphone and camera. This can only be done from within the conventional 2D

webpage before immersive mode begins. Care must be taken by every website designer to handle those steps sequentially.

Every WebXR site ends up being a small walled garden of its own. Even though they all sit inside the same browser, only one can be in immersive mode at a time. This isn't a property of the WebXR canvas, but an intentional feature of the web browser architecture that keeps websites apart from each other. WebXR developers must build an entire stack that enjoys very few of the benefits that web browsing offers, and those experiences invariably end up with most of the same walled-garden friction as the closed platforms.

Finally, here comes the bad part.

A browser cannot build one scene that many companies own

Web browsers are architected from the ground up to maintain exceedingly strong isolation between multiple origins, or websites. The only provision it has to allow one website to embed within another is an iframe, which was implemented on top of earlier foundations using sandboxing and site isolation to keep cross-site frames separated and contained. This does not allow for content from several origins to be merged into one scene. Unfortunately, iframes don't extend into WebXR, meaning there is no facility to add one WebXR canvas inside another – not that you'd legitimately want to.

That's not to say a webpage can't be composed of documents that come from several origins. It's very common to see scripts, stylesheets, maps, video, and payment forms from different origins composed together on a single page. But the author of the page must decide in advance what gets composed and must have access determined by the external server's CORS configuration.

This decision – embedded at the deepest architectural level of every web browser – to delegate all decisions of multi-origin composition to the main webpage loaded into a tab is entirely antithetical to how the metaverse needs to operate. One webpage operator cannot possibly have predetermined knowledge of every single origin, location, and service you might encounter as you travel anywhere in the physical or virtual world in order to preconfigure CORS. Nor should it.

The act of browsing the metaverse must be handled intrinsically by the browser itself, not by a single page loaded into a browser tab. This is because metaverse content is fundamentally multi-origin, potentially hundreds at once, where all origins have equitable access to the scene and resources of the browser. WebXR is not getting in the way of this. It's just an innocent bystander that draws pixels. The web browser architecture itself is what prevents web browsers from properly browsing the metaverse.[75]

Why companies built closed systems instead of waiting

A majority of Fortune 500 companies have XR products on their roadmap and have been actively developing products for many years. Their options for developing XR products are fairly limited and include WebXR, HoloLens, gaming engines like Unity, Unreal, and Godot, or proprietary third-party applications. These products are common for training, prototyping, and proof of concept.

Very few have made it into a production setting and even fewer have reached the consumer. This is primarily because these solutions are built on top of closed platforms. Even WebXR, which is open, but most limited in its capabilities, requires specific devices like a Quest 3 VR headset to function well. By and large, none of these solutions can be deployed across the devices, vendors, and outside parties that their day-to-day operations rely on.[76]

They built closed platforms because the open platforms they needed did not exist and the alternative was to ship nothing at all. Inside their own walls they supplied every missing piece, from native rendering and real-time performance to precise tracking, identity, and direct access to the device. What they built renders beautifully, but only in controlled training labs operated by employees in white coats. These systems aren't ready for factory floors, fast food restaurants, or anywhere outdoors.

Proprietary, closed platform training applications are not the same as browsing. They are built the way games are built using game engines like Unity, Unreal, and Godot. Such tools are made for shipping a title rather than for browsing a world, and content built in one engine and for one device will not run in any other configuration. Every device a

company wants to reach needs its own build. Updates require first the developer to push the product and then the user to install it.[77] Few, if any of the intrinsic features of the metaverse are built in, such as the ability to discover content and services based on proximity.

The open web and game engines together hold just about everything the metaverse needs. The web brings device independence, lack of friction, open standards, and self-hosted content. But the web cannot deliver performance, sufficient memory, local storage, threading, multi-origin scenes, and direct access to cameras and sensors that demanding three-dimensional scenes consume. Native game engines are the exact complement.

In the absence of both halves of the equation in one product, enterprises were forced to choose the one half they could control and ship – settling for game engines over open standards. The result is that each of them now have their own proprietary solution that lacks interoperability with one another.

Large organizations and robots cannot work inside closed systems

The irony of all this is that the very enterprises with the deepest need for spatial computing technology, such as airports, hospitals, banks, and municipalities, are exactly the ones that absolutely cannot operate inside walled gardens. They cannot put their operational data on a platform they do not control, and they will not stake critical operations on a company that might change its terms or walk away.

The industry we have today is fragmented, locked in, and carrying platform risk that grows with every new deployment. When it was only people wearing headsets, the fragmentation was tolerable. Now that robots, autonomous vehicles, and AI agents need to share spaces with humans, every proprietary wall becomes an expensive deployment and safety barrier. The net result is often that XR projects get stuck in a perpetual proof-of-concept mode, otherwise known as product purgatory.

The underlying structure must be replaced, and trust must change

The way out of this mess is quite straightforward. Develop a metaverse browser that retains all the browsing properties of a web browser, but that is architected to meet the requirements of a spatial metaverse. This is easier said than done, but it's not an impossible task.

Work starts at the lowest layer of architecture, replacing a single-origin DOM (Document Object Model) with a multi-origin SOM (Scene Object Model) capable of accepting content from many domains at once, yet making sure they are each sufficiently protected from each other. Content and services must be able to flow in and out of the SOM without manual installation as you move around, based on proximity rather than when you click a hyperlink or type a URL. The browser must know where you are precisely enough to attach virtual content to real content with centimeter (or sometimes millimeter) accuracy. And you must be able to carry your identity from one place to the next instead of having to recreate it at every threshold.

Allowing many origins to co-mingle in a single scene might sound like giving up the protection that origin isolation provides. But it doesn't have to be. Each service can still run sealed off from the rest, unable to read private data that another service beside it is processing, or touch anything on the device it was not granted permission to access.

In web browsers, origin security is tied exclusively to the domain specified in the URL. In a modern, multi-origin architecture, security is better derived from signatures held in certificates that travel alongside the content, which the browser can easily verify. This puts trust in a verified publisher rather than a host name.

Fortunately, these are not just far-off ideas. All these concepts already exist as working technology today.



Chapter 10

Opening Up Is the Bigger Prize

> ***"There is an almost mindless debate about the Internet versus online services, the Web versus AOL, as if there's some inherent tension between them."***
>
> **— Steve Case**
>
> *co-founder and CEO of AOL*
> *1996*

Closed, spatial platforms are repeating AOL's wager

Every company building a closed spatial platform today is making a bet that, this time, the walled garden beats the open web. That bet has been placed before, by companies larger and better positioned than anyone in spatial computing today. In the 1990s, before the open web won, most people's first experience of going online didn't happen on the open internet at all. It happened inside a walled garden – America Online, CompuServe, or Prodigy. You did not browse the web. You browsed one of these company's sliver of the web. They were convenient, curated, safe, completely closed with a wall around it, and one company guarding the only gate.

Prior to 1992, an AOL user could not send an email to someone on CompuServe, because the two services' mail systems were incompatible. Later, you could send limited-sized messages between the two for a minor fee.[78] Spatial computing is walking back into the same trap. However, the cause of AOL's downfall is the strongest reason for optimism about where spatial computing goes next.

The early walled gardens weren't entirely without merit at the time. The open internet of that era could not deliver a smooth, safe browsing experience to a mainstream audience, so companies built closed platforms that could. The bargain exchanged convenience for operating within someone else's walls. Spatial computing is following the same playbook and for the same structural reason. The open web can render 3D and stream in real time, but does so only within a single document from a single origin. There is no way to compose one continuous space with many independent providers, to discover what is around you as you move, or to carry your identity from one place to the next. Closed platforms can offer some version of those features today. Few of these platforms last. Not because their technology didn't work, but primarily because a closed system cannot compete with the reach of the open web.

The expense of running a large closed system at scale takes a hefty toll on the operator. Building large platforms is extremely expensive, ranging from many millions and even into billions of dollars. The justification for that type of expenditure stems from companies like AOL and Apple. Own the platform, own the store, own the developer relationship, and the investment will eventually drive a highly profitable business. Boardrooms often view the walled-garden approach as a way to protect their investment.

These companies find themselves balancing two opposing forces – creating demand and enclosing it at the same time. Without their platforms the market will not exist. But without an open ecosystem, users – especially enterprises – often encounter too much friction to justify adoption. This is because enterprise operations span a variety of devices, vendors, contractors, and customers, so they cannot operate inside one company's walls. Consumers must be able to find enough

content and utility from new devices to justify owning them, let alone replacing their existing devices.

AOL lost because the open web offered reach and ownership together

AOL did not fade because it was poorly run. At its peak in 2000 it had more than 26 million subscribers in the United States alone, more than 40 percent of American internet users at the time, and a market capitalization of roughly $163 billion.[79] It faded because the open web offered two things it could not compete with.

The first was universal reach. Any person could visit any site, on any device, without permission from a gatekeeper, and anyone could publish without asking. The second, and more important, was ownership. A business could build its own presence on its own domain and its own servers, owning its customers and its data, rather than operating within AOL at their discretion. On the open web it owned what it built, and nobody could take it away.

Both properties compounded, because every new site made the web more valuable and every user who came online made publishing more worthwhile, a network effect no single walled garden could compete with. Reach is what made the web popular, and ownership is what made it safe to build a company on. Note that the web did not attract the world's businesses and then open up. Rather, the world's businesses were attracted because it was open. Operating on open infrastructure, as it turns out, is non-negotiable for virtually every business on the planet.

This leaves businesses with a bit of a conundrum because the only systems that can do spatial today are closed ones, but closed systems cannot offer ownership. Like AOL, ownership is the business model of closed platforms, leaving businesses little choice, since, for the time being, there are no viable open solutions.

With the largest audience online, the best-known brand in the business, and distribution no one could match, AOL was in prime position to become one of the open web's biggest winners. It chose to fortify and defend its walls instead. That did not end well for them. The open metaverse is in its infancy now. If history has any say in the

matter, it would behoove today's walled-garden companies to take down their walls and embrace it.

Opening up is the larger commercial prize, not the noble sacrifice

The open web created vastly more value for vastly more companies than the walled gardens it replaced. It did not divide AOL's market among competitors. It created an economy orders of magnitude larger than AOL ever touched. Alphabet, Amazon, and Meta, all three of them built on the open web that replaced it, are together worth more than $8 trillion, nearly fifty times AOL's peak.[80] The same dynamic is at play in the metaverse, because an open spatial layer expands the market for every device maker, every enterprise, every developer, and every chipmaker. Openness is not an altruistic choice that sacrifices profit. It is the commercially larger choice.

Meta has absorbed roughly $83.6 billion in operating losses on Reality Labs since 2020, nineteen billion of it in 2025 alone.[81] It is the largest bet any company has ever placed on spatial computing. From a purely engineering perspective, it produced some spectacular technology. However, from an economic perspective – let's just say the road to profitability is a long way over the horizon. Content sold on its store has earned somewhere above two billion dollars in total since 2019, a figure the company gave in 2023 and then gave again, unchanged, eighteen months later.[82] A single gaming console platform turns over more than thirty billion dollars in one year.[83]

Starting platforms from scratch is never easy because users are attracted to content. As a result, the platform had to fund content development itself. When the funding was pulled back, the projects went with it. Meta's own chief technology officer said the gravy train of paid-for content had come to a stop. It had won over enthusiasts without ever materially expanding the market.[84]

A closed ecosystem also must convince developers to learn a toolchain that works in only one place, which is difficult to justify when the reward is a single device's audience. It also poses a liability in the event the ecosystem doesn't survive. Open ecosystems, in contrast, have a completely inverted risk versus reward proposition for

developers. What a developer learns carries everywhere and their work will be protected against any individual device or platform failure.

Meta's losses reflect closed-system friction

The glass half empty observer might look at Meta's failures and conclude that the world has no use for spatial devices. The more likely explanation is that there was too much friction, too much risk, and not enough opportunity for expansion in and beyond the closed ecosystem that Meta purpose built for a very niche audience.

That lesson should not be ignored by manufacturers attempting to introduce spatial hardware devices in today's market. These companies have much more to gain by building them to run on the open metaverse. A device maker selling smart glasses is capped today by enterprises and universities that will not standardize on hardware tied to one company's platform. Open infrastructure is what turns them into volume customers, because it lets them commit to the category without committing to a single vendor. The same holds for the app and tool companies, whose addressable market currently stops at the edge of whichever garden they are stuck inside. Open standards do not shrink any of these businesses. On the contrary, removing walls from a walled garden simultaneously removes the boundaries that prevent companies from growing well beyond them.

As a nice bonus, the companies that help build the foundation get something valuable in return. They have an opportunity to shape standards while they are still being written instead of implementing someone else's decisions years later. These companies will enjoy the advantage of being market leaders. Some of the biggest winners in the web space were the companies that stepped in early to provide hosting, authoring tools, browsers, and search – the layers that grew on top of the standard. Every one of those roles is unclaimed in spatial computing today.



Chapter 11

Enterprises Need Infrastructure They Control

> ***"Control your destiny or someone else will."***
>
> **— Jack Welch**
> *former CEO of General Electric*

Enterprises learned to say no by watching platforms disappear

In today's spatial computing landscape, enterprise decisions frequently hinge on one overriding question about the platform underneath it. Not *does it work?* Not *what does it cost?* The question is, w*hat happens if the vendor shuts it down?* We've seen that exact scenario happen time and time again. Until that question is no longer necessary to ask, almost every potential product will remain in a state of perpetual proof-of-concept.

HoloLens, Microsoft's pioneering mixed-reality headset that many enterprises built real products on, was wound down, leaving organizations stranded mid-rollout, and their investments orphaned.[85] 8th Wall, a widely used web-AR platform, was acquired and then

deprecated.[86] AltspaceVR, a social-VR space people had built communities and events inside, was shut down effectively overnight.[87] Ready Player Me, an avatar platform that creators and companies built on top of, was acquired by Netflix and turned into a closed platform product.[88]

These companies exemplify the risk of building on a platform controlled by another company. Their strategy shifts, an acquisition changes priorities, a product line stops making sense on someone else's balance sheet, and your mission-critical deployment evaporates, on their schedule, not yours.

The lesson learned: when you build on a proprietary platform, the platform owner can end support with little or no warning, and everything you invested goes with it. Your time, assets, programs, and skills are deprecated overnight. In the best case, you've lost a lot of money. In the worst case, your company's business operations are entirely disrupted.

Integration costs keep deployments trapped in pilots

When Microsoft, Intel and IoT Analytics surveyed five hundred manufacturers about digital twins, twenty-eight per cent had begun a rollout and only one in seven of those had it fully deployed in their plants, while another forty-five per cent were still in development or proof of concept.[89] The obstacle they named most often was integration, three-quarters of them, followed by the time it takes to build one and the overall complexity of the system.

Those findings don't suggest a failure in the technology. They only point to the costs involved in connecting systems that were never designed to work together. Microsoft's report concluded that the industrial metaverse was not yet a priority, which is a fair description of a market where the work is possible, but the plumbing makes it unaffordable. That survey was taken in 2022, and in the years since, the number of vendors has grown, the number of devices has grown, but nothing has materially improved to connect any of them.[90]

There is one specific use case that has gained a foothold in the enterprise space – training. This is because a training system can sit in its own corner, installed on its own specialized equipment, and be

replaced or discarded without touching anything the business actually runs on. Anything that might touch live operations, private company data, or paying customers is precluded due to the aforementioned risks. Which is why a decade of enterprise XR has produced an abundance of training pilots, but very little else.

Five enterprise requirements decide whether a deployment ever ships

The entire world already runs on web infrastructure. Communications, logistics, inventory, payroll, sales, collaboration, customer records, and the AI tools now stacked on top of all of them are delivered through a browser to any device you happen to be using. Every business, every airport, every hospital, every bank, and the laptop sitting in your house depends on that same open web, precisely because no one owns it and no one can take it away.

As computing moves off the screen and into physical space, why would the requirements suddenly change? Metaverse infrastructure will carry operations at least as critical as anything a website handles today, and probably more critical, because it acts in the real world in real time.

Despite the sparse number of deployments at the enterprise level that rely on metaverse infrastructure today, the amount of interest and investment is incredibly large. Enterprise executives and IT departments across the spectrum are actively looking for spatial solutions that solve their company's particular needs.

This dichotomy has walled-garden service providers, who dominate spatial computing today, wondering why adoption of metaverse infrastructure is so slow. The reasons are no different than what businesses have required over the past two decades when transitioning to web infrastructure. These are the non-negotiables, any one of which can permanently prevent a project from reaching deployment.

Device agnosticism. A deployment must work across devices and manufacturers rather than locking the organization to one vendor's headset or glasses. Device manufacturers are hardest hit by this requirement, which shows up as a ceiling on their sales. Enterprises simply will not tie themselves to a single product or manufacturer, as

the HoloLens veterans will attest. For consumer-facing businesses, there is another very valid reason for this requirement. Businesses know that they can't control which devices a customer will be using when they walk through the door. The solutions they choose absolutely must work with all consumer devices.

Data ownership. An organization's data, often sensitive and sometimes regulated, must live on the organization's own infrastructure under its own control. In healthcare, finance, education, and defense that is not merely a preference, it is a legal requirement. Hospitals can't hand patient records to a company that hasn't signed on to be legally responsible for them, universities can only share student records with services they directly control, banks answer for every vendor that touches their customers' data, and government agencies and defense contractors are limited to cloud services the government has approved.[91]

Frictionless access. Content and services must be accessible without having to install a separate application on every employee's device. The alternative is a separate app for each vendor with its own login, update cycle, and support queue, handed to thousands of employees who use a variety of devices. That becomes a logistical nightmare for any medium to large organization. It is far more manageable to maintain software applications on a server that employees can access with browser-based clients. When maintenance is needed, a software system can be deployed in one central location, and everyone is updated at the same time without needing to change anything on their devices.

Open developer ecosystem and open tooling. Anyone must be able to build and deploy metaverse infrastructure software and hardware without a gatekeeper's approval, so long as it conforms to the standards. This applies equally to content. On closed platforms, the operator provides a limited set of tools, typically based on those that generate the most revenue for the operator. Companies that require other tools and services have no option but to wait for the operator to provide them. On an open foundation the organization can build it in-house or source it from a third party in the same way the open web works now.

Interoperability. Content and services must work seamlessly between systems and vendors rather than locking the organization to

one specific supplier. Inside a business, operations rely on sites, contractors, and suppliers who use a variety of software and equipment. With today's web infrastructure, these components are all designed to operate on one single web standard. Whereas today's closed ecosystem approach creates proprietary systems built by different companies with different standards that are not interoperable. An open metaverse provides interoperability standards that all vendors can create for, ensuring that data created in one service can be easily transferred to another.

The web provides a proven model for every requirement

The net effect of companies building their operations on closed systems is a pattern that plays out across the industry. Someone inside the organization understands the value immediately – smoother operations, safer floors, faster training, a better experience for the customers walking through their doors – so they build a prototype. The prototype goes up the chain. When it reaches IT, security, and executives, the questions change from functionality to viability. *Who owns the data, and where does that data reside? What happens to this investment if the vendor pivots or shuts the product down?*

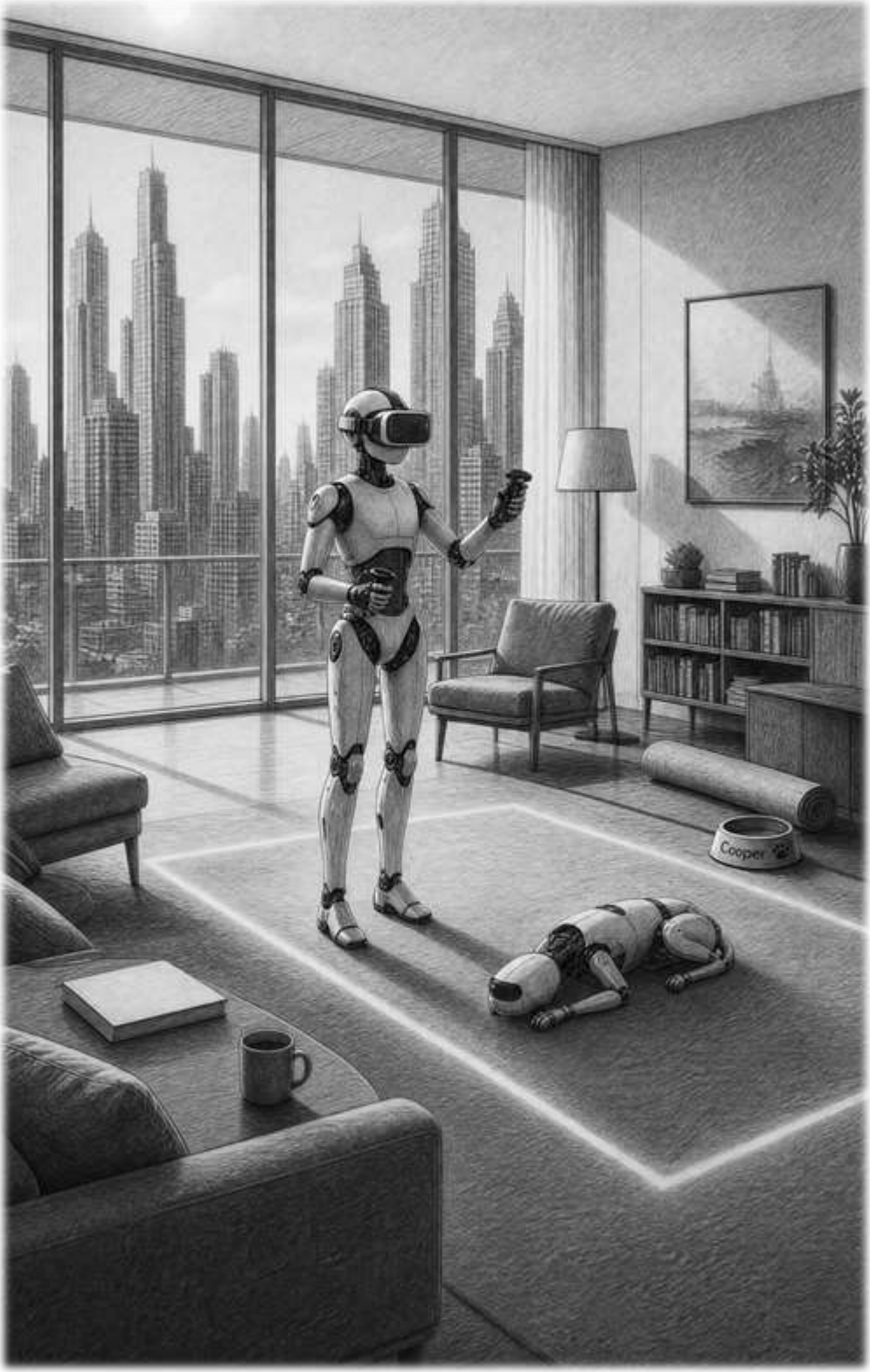
Unless the answers are that the organization owns its data and cannot be stranded, the project halts. Project after project ends up in product purgatory. Endless pilots dazzle and then die, never reaching production. Not because the technology is bad, but because no responsible executive will bet the operation on a single company that can terminate their product overnight.

Every one of the non-negotiable requirements is something the web already guarantees – for documents and information. Any device can view any webpage. You can host your own content on your own server. You browse to a page without installing anything. Anyone can publish without permission. And everything interoperates, on a web that outlives any single company because it belongs to no one. The reason the web became critical infrastructure that most institutions on earth rely on – banks, governments, hospitals, airlines – is precisely because it satisfies these conditions.

Metaverse infrastructure will not earn the same stature, and therefore will not escape proof-of-concept purgatory, until it satisfies them too. Enterprises are not asking for something new or unreasonable. They are asking for the web's guarantees, evolved for physical locations. The open metaverse is that solution.

Part 4

Spatial Computing Needs a New Foundation


Cooper

Chapter 12

Metaverse Browsers Are Still Browsers

> ***"The browser is the lens through which people see and touch the Internet."***
>
> **— Mitchell Baker**
> *co-founder of Mozilla*

Four properties make a browser a browser

The average person spends more than six and a half hours online every day, roughly 40 percent of their waking hours.[92] A good share of that time is spent using a web browser. The Cambridge Dictionary defines a web browser as, "a computer program that makes it possible for you to read or see information on the internet."[93] That might be the understatement of the century. It omitted many concepts such as retrieving content other people publish on their own server and the extensive use of open standards. One could fill volumes trying to explain every possible feature packed into a piece of software estimated to contain nearly 40 million lines of code.[94] Here are four properties

that we believe make a browser distinct from every other kind of software, and which are, at their essence, what make a browser a browser.

Trust without knowledge. You can safely open a website you've never heard of, run by an organization you know nothing about, without worrying about whether the content will harm your device. The browser runs every site in a sealed box, so a site cannot alter your device, read your files, reach into another site you have open, or gain access to information about your browser or sites you may have visited in the past. That protection is what lets average people use the web. Without it, only experts would risk visiting a site they don't recognize, and the web would never have reached billions of people.

No installation. Besides the browser itself, nothing else must be installed before a website works. You never see a setup screen, and no software remains on your device after you leave a webpage, unless you allow it. You arrive, the content runs, and when you leave it is gone. Every additional step introduces friction, and the browser eliminates the steps of download and installation entirely.

Any device. The same webpage loads on a laptop, a phone, a tablet, or a television. Its publisher doesn't know which device you are using and doesn't need to since it publishes its content based on a standard, rather than a particular product. The browser is responsible for displaying content based on the same standard, so a webpage written once should work across the board.

Content kept separate from its display. A publisher describes what content should appear alongside a description of render intent, but ultimately the browser decides how to show it. This is true for large monitors, phones turned sideways, and screen readers for someone who cannot see. That separation is why nearly every business, school, and home can afford to create a website. The browser allowed people who did not know how to program computers to publish information on the internet.

Omit any one of these properties and you no longer have a browser – at least one that people would be willing to use. Roblox and Fortnite are good examples. Each one must be installed, they only work on devices and platforms for which they have compiled programs,

games are not able to be hosted on your own server, and you can only play games that are subject to the approval of the platform. Both fail all four tests, which is why nobody calls them browsers.

Metaverse browsers apply the same four properties to three-dimensional content

A metaverse browser must be able to deliver all four properties, or it won't adequately serve the same browsing function as a web browser. It's not enough to just add a third dimension to the graphics. Browsing the metaverse adds extra demands that browsing the web doesn't have to contend with.

Trust without knowledge. All the locations you encounter still run services in their own sealed box, the same protection a website gets. The difference is scale. A web browser only has to manage one site at a time per tab, while a metaverse browser may be running code from dozens or hundreds of locations concurrently as you move through a street or a building.

No installation. Nothing must be installed to interact with a location. The difference is frequency. Someone browsing the web opens a new page every few minutes or so, while someone browsing the metaverse will encounter new services every few steps.

Any device. All the content you encounter must still be viewable on any device. The difference is variety. A webpage mostly has to fit different screen sizes, while metaverse content is experienced immersively on devices that include mobile phones, smart glasses, and VR headsets.

Content kept separate from its display. The content the browser displays is still independently published. The difference is its location. Every object in the metaverse has a three-dimensional transform relative to its parent object, which must be synchronized with your own position and the position of other objects in the real and virtual world.

Spatial content is attached to the world, not floating in front of you

A webpage or a video floating in a flat window placed in your field of view is not spatial content. That's just 2D content presented in a

window on a 3D display. Flat material is optimally rendered on a surface you can touch or point at precisely and provide input with a keyboard. Without that surface the precision is gone. Phones and laptops are still the right tools for reading, typing, and most 2D work. It will be very difficult for smart glasses to compete with your phone on the basis of two-dimensional content.

Spatial content, on the other hand, has a position in the world instead of a position on a screen. While your phone can be used to view that content using its camera and display, it lacks immersion. Smart glasses can display that information in a convincingly realistic way that flat surface displays cannot. For this type of experience, smart glasses and other spatial devices have the upper hand.

Virtual spatial content blended with the real world poses its own set of challenges. The browser must know precisely where you are in relation to your surroundings or else it won't know where to place the virtual content. Furthermore, that content must remain in place as you move around it. Even trickier, virtual objects can get obscured by real objects that sit in front of them. To do this, a browser must be able to detect real-world objects and be able to determine their exact shape, size, and position.

Spatial content that is hosted by the locations you encounter must also be merged with the content that you bring with you into the metaverse. This content is called overlays, and includes such things as navigation, store reviews, and virtual pets that travel with you as you move through the metaverse. The metaverse browser must manage the task of merging all this content together into one unified scene.

Metaverse browsers need three things web browsers never had to provide

Metaverse browsers aren't just software programs that display three-dimensional content. They must provide all the essential properties that make them browsers. But they must also possess a few new properties of their own. This is because the nature of browsing the metaverse is very different from browsing the web. You cause a web browser to load pages at whim – either by following a hyperlink, performing a search, or typing a webpage URL whenever you feel like

it. Conversely, metaverse browsers load content based on your proximity to it. Every object and service is mapped to a three-dimensional coordinate, which you access merely by moving yourself near it – both in the real and virtual world. Because of this fundamental difference, metaverse browsers require three things web browsers never had to provide.

Finding places by proximity. Websites have no coordinate position relative to one another. They are referenced only by domain name, which is used to obtain the IP address of the server on which they are hosted. The content returned is positioned relative to the origin of the page – the upper left corner of the browser tab in which the page sits. While spatial content is also accessed in the same manner from servers over the internet, every object returned has a 3D coordinate transformation that defines its exact location in a map. This map can exist only for one location, or it can be a map that is as big as the universe. The browser synchronizes your real or virtual location with its map, making it possible to automatically load content near you.

Staying live. While some webpages provide live content, the overwhelming majority of the web consists of static pages that update only when they load. The opposite is true in the metaverse because locations, by their nature, are constantly evolving. Your mere presence in the metaverse changes the space around you. Co-presence adds an additional dimension because everyone needs to have a consistent view of the spaces they share. This includes not only the content and services provided by each location, but also the movement and voices of the avatars passing through them.

Many owners in one view. Browsing the web is done one page at a time, where each webpage has complete authority to decide how content is arranged in its own browsing tab. This assumption fails when browsing the metaverse because all content from the many independently operated locations near you is composited into one three-dimensional scene. No single content provider is considered more important than another, so they must all share the scene equitably.

A shopping mall provides a good example of all three of these metaverse browsing properties in play. Because it is so large, the stores

are not all loaded into the browser at once. As you walk through the mall, the stores you can see, as well as the content they present, are continually being updated based on your interactions and those of the people around you. What you see and hear is constantly evolving. The content from all the stores in your browser along with the avatars of the people around you are mapped into one shared scene represented by your device. It's worth mentioning that the browser does so safely and securely so that no store can peer into the content and services of another if they choose to make it private.

People and robots browse the same published world

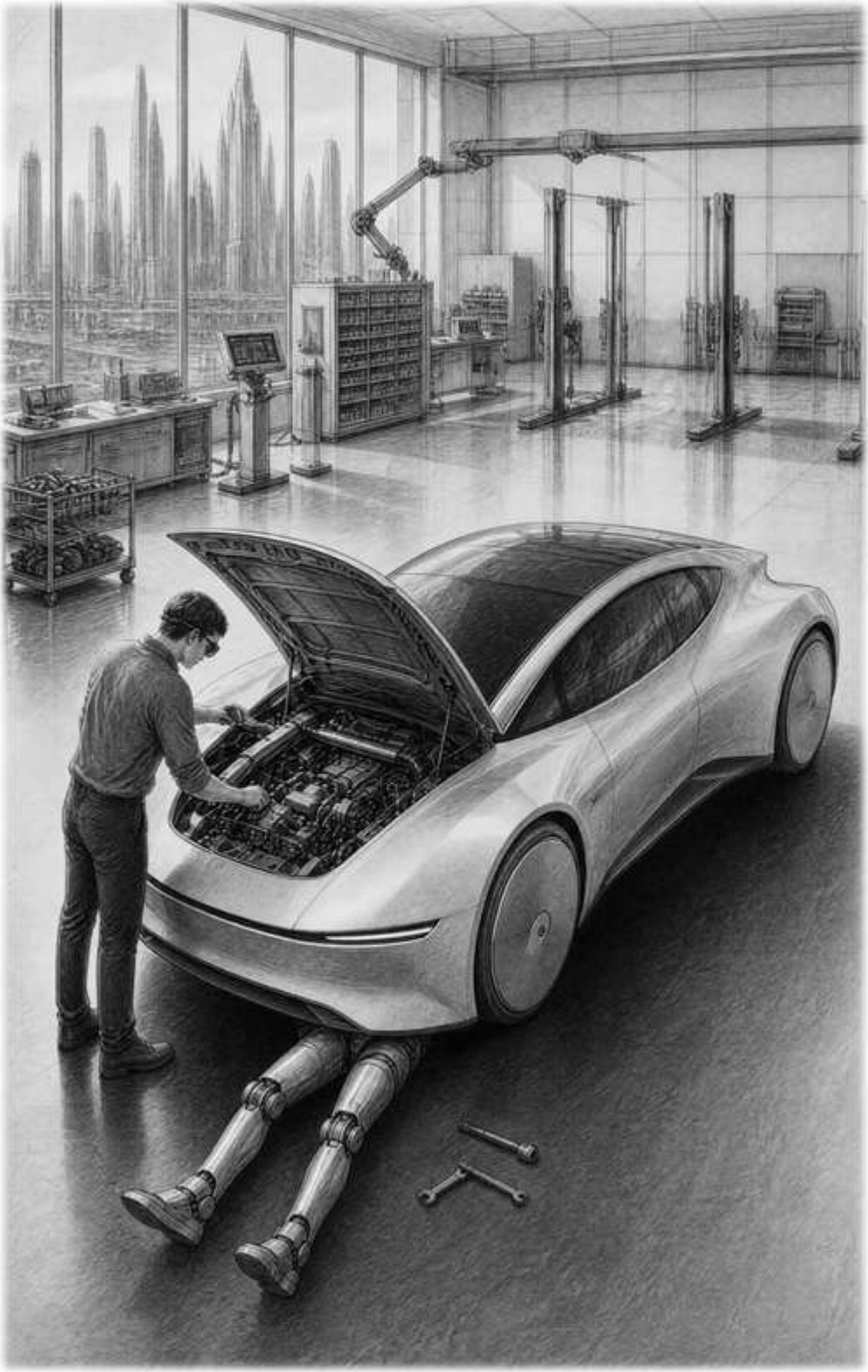
Robots, autonomous vehicles, and physical AI agents need to browse places for the same reasons people do. Each of them operates within a location and needs to know where it is, what's around it, and which services are available for it to use. A metaverse browser provides this via the same mechanism that it provides to humans.

The open metaverse provides benefits to all these non-human entities that us humans can't take advantage of – or at least not nearly as easily and quickly. Humans must use our own input devices – eyes, ears, fingers – to interpret the world around us. We adapted our world to provide clear signals that we can easily interpret, such as stop lights, signs, and buttons.

Expecting robots and autonomous vehicles to rely on our signs is not the most efficient mechanism for them. They must use cameras to replace our eyes, microphones to replace our ears, and fingers to push our buttons. And people need to build incredibly smart software to interpret signals from these devices – which doesn't always work out well for nearby humans.

It would be far better if robots and machines could connect directly to metaverse infrastructure to receive safety signals from a piece of equipment or communicate directly with an elevator. Or even to connect with humans. A robot present at a location could signal people who can't even see or hear that robot by injecting a virtual safety fence into a location, thereby warning users wearing smart glasses to avoid an area.

A published digital twin that adheres to the standards of the metaverse works the same for everyone and everything that enters that location. The people walking through it, the robots working in it, and the AI acting for either of them all read from and interact with one published description, each with only the permissions it has been given. In that way, the metaverse acts as the common infrastructure that all these entities communicate through, eliminating the need for a thousand different communication protocols between each of them communicating separately.



Chapter 13

A Shared Engine Enables the Metaverse

> ***"Talk is cheap. Show me the code."***
>
> **— Linus Torvalds**
>
> *creator of Linux*

Common standards and shared engines made the web portable

Every web browser has an engine inside it, though very few people can tell you what a browser engine is or even any of their names – Chrome runs on Blink, Safari runs on WebKit, and Firefox runs on Gecko. Web browser engines process the HTML, CSS, and JavaScript files that comprise a website into the webpage you see and use. They also enforce the safety boundaries between sites and your device, so a page you opened by accident cannot access your bank account in the next tab or the documents on your hard drive. The browser is the entirety of the program you install and interface with. The engine does the work of rendering the webpage.

Open web standards developed by groups such as the W3C, WHATWG, and Ecma International define what a webpage is and how it should behave; engines implement those standards. Because the standards are common and the major engines are shared across many products, a company can confidently publish one website instead of writing different versions for every combination of browser and device. Chrome, Edge, Opera, Brave, and Samsung's mobile browser all use Blink while Safari uses WebKit, but they all read the same HTML, CSS, and JavaScript.[95]

The same principle holds true for all types of devices, even VR headsets. Instead of writing a new web browser engine from scratch, Meta adapted its Meta Quest Browser from Chromium, the open-source web browser behind Chrome, whose engine is Blink.[96] An ordinary standards-based website can therefore open on a Quest without its publisher building a separate Quest application.

Metaverse browsing needs a shared architecture

Metaverse browsing needs a metaverse browser engine of its own. A webpage has one top-level document whose author decides what appears in it. But a physical place has no single author. An airport, for instance, may contain services from the airport facility operator, several airlines that operate within its terminals, government agencies, restaurants, and shops, all accessed by the employees and travelers walking through it. All these entities within the airport need to contribute to a single composed view without giving any one of them control over the others.

In the absence of a common metaverse browser engine based on shared standards, each company has had to build some limited version of that machinery for itself, on its own devices, and in its own format. Content made for one system will not run on another's. Each new platform begins with an empty catalog. This creates an incredible amount of redundancy. A company that creates a 3D model of its product, store, or factory must bundle it inside of an application, distribute that application, and persuade people to install it. The web ended that problem for informational content a long time ago, and until recently, the same solution was not available for spatial content.

Sneeze can run wherever browsing happens

Sneeze, a new metaverse browser engine, processes spatial assets and services into a live three-dimensional scene in the same way that Blink turns website code into a two-dimensional document. A publisher can host its spatial content and services online for any browser or device using Sneeze to discover, stream, and render them into the scene without requiring a person to install a separate application. Sneeze does not replace the web or require everyone to buy a new kind of device, and it can be deployed in three ways.

The first is to embed Sneeze directly into an existing web browser alongside Blink, WebKit, and Gecko. The browser continues to open existing websites in its web browser engine and metaverse locations – known as spatial fabrics – in Sneeze, its metaverse browser engine. Chrome, Safari, and other browsers can simply add metaverse browsing without giving up anything they already do.

The second is to embed Sneeze directly on a device. Smart glasses or a headset can boot into a thin metaverse browser that uses Sneeze to browse the metaverse without having to manually launch applications. Devices that use this approach become a commodity, requiring no multi-billion-dollar investment to create and maintain a proprietary ecosystem around their hardware product. Like other accessories such as watches or mobile phones (within one of the two major platforms), these devices will compete on performance and style since they are purpose-built to provide frictionless access to the open metaverse.

The third is to deploy Sneeze within a standalone metaverse browser. RP1 adapted Artemis directly from Rubidium, a publicly available open-source browser that embeds Sneeze. Like Chromium for web browsers, anyone can use Rubidium as a starting point for building a metaverse browser of their own.[97] In theory, a standalone metaverse browser could place Blink or WebKit beside Sneeze supporting both webpages and spatial fabrics in one product, but that functionality does not yet exist in Rubidium.

Using one engine does not make every product the same. Device makers and browser companies still control their hardware, interface, services, and brand. They share the machinery underneath, just as

Chrome, Edge, Opera, and Brave all use Blink, but remain different products with unique feature sets.

Open code lets the industry share one engine

Sean Mann and Dean Abramson are two of the founders of RP1, which builds commercial technologies and products that take full advantage of the benefits made possible with metaverse browsers, such as RP1, their universal, public spatial fabric.

Sneeze has been publicly available as open-source software since June 2026 under the Apache 2.0 license.[98] It is currently in early alpha and under active development. Tim Bajarin, reviewing the Augmented World Expo for Forbes, described it as, "Blink and WebKit reimagined for spatial computing."[99] Rubidium has been publicly available as open-source software since September 2026 under the Apache 2.0 license.[100] It is currently in early alpha and under active development.

Sneeze did not start as a native engine. The first metaverse browser we built ran on WebXR, on top of the web. It worked well enough to show exactly where the web stops, so we started over and built the engine natively. Part of the way through, it was clear that the engine could not belong to any one company, including ours. An engine that one company controls is just another walled garden.

Sean, who serves on the board of the Metaverse Standards Forum, and Dean, the inventor and architect of Rubidium and Sneeze, brought these open-source projects to Neil Trevett, president of the Forum and the Khronos Group and vice president of developer ecosystems at NVIDIA.

The Open Metaverse Browser Initiative launched in March 2026 as the Forum's first Initiative. The open-source projects sit outside the Forum so that anyone can contribute without becoming a member. A business council connects the projects back to the Forum.[101]

The Metaverse Standards Forum does not write standards. It serves as a central meeting point for standards bodies and other organizations involved in standards, including the W3C, NIST, the Khronos Group, the Open Geospatial Consortium, and Open AR Cloud. Its dedicated focus is metaverse interoperability. Both Sneeze and Rubidium rely on standards from several of those organizations,

which makes the Forum the natural place to bring the browser and engine together. Those open standards include glTF for 3D assets, WebAssembly for portable sandboxed programs, and OpenXR for XR devices.

The metaverse browser and engine both must be open if they are to become common infrastructure. Their Apache open-source licenses mean that a device maker does not have to build its product around software a competitor controls, and an enterprise does not have to rely on code that one vendor can withdraw.

The Apache open-source licenses also mean that the underlying code will never go away. If one contributor stops working on it, another can continue. Device makers can embed it, browser companies can adopt it, enterprises can inspect it, and developers can build for it without asking permission from the company that invented it.

Metaverse browsing now has the common engine it was missing. The web browsing paradigm applied to metaverse browsing allows individuals and companies to focus on publishing spatial content and services on their own servers without needing to bundle them into device-installed applications. You publish your content; the metaverse browser handles its delivery and display. The open code that underpins this metaverse infrastructure allows the industry to standardize on it without handing over control to any one company.



Chapter 14

Metaverse Browsers Make Spatial Publishing Easy

> ***"We are stuck with technology when what we really want is just stuff that works."***
>
> **— Douglas Adams**
> *author of The Hitchhiker's Guide to the Galaxy*

Game engines bundle the renderer, making shared scenes difficult

Most people who need to create rich 3D content today choose to build on top of a game engine such as Unity, Unreal, or Godot. Unity reports that about 69 percent of the top 1,000 mobile games are made with its engine alone.[102] Game engines produce sophisticated, compiled, native applications. The renderer, the physics, the networking, and the content are all packaged into one program for each device that users must download, install, and run. A different version of the software must be recompiled for each type of device or platform.

WebXR, the standard that brings immersive 3D to web browsers, works much differently. It renders pixels added to a canvas element on a webpage and supports immersive mode to VR headset. But the website must supply everything else including the 3D engine, networking, and avatars. Because WebXR runs in a browser, users don't have to download or install any software, but they struggle to deliver demanding 3D at the quality that game engines can deliver. Browsers can run only one immersive session at a time, so two WebXR sites cannot share a scene any more than a Unity, Unreal, or Godot experience can.[103]

Today's landscape of existing 3D solutions is the cause of the current fragmentation in spatial computing. Every option available today makes you entirely build your own spatial computing stack directly into your application, making it incredibly costly, and difficult to impossible to combine with any other standalone product. Because each application carries its own renderer, it is not possible for two applications to share one scene, which is one of the main reasons why we need a new metaverse browser engine.

The DOM is why a webpage is easy to publish

A web browser organizes every page into one tree called the Document Object Model, or DOM. Each shape, paragraph, image, and button on the page is an element in that tree, and the browser's layout, scripts, and rendering all work from it. The W3C published the DOM as a standard in October 1998, and engineers from Netscape and Microsoft, the two fiercest browser rivals of the day, were among its editors.[104] Because every browser follows the same standard, a webpage that conforms to that standard works in all of them.

The author of a page owns its DOM exclusively. The author describes the information that the document contains along with its intent for how it should be rendered, but the browser ultimately decides how to draw it. That is one of the key properties of browsers – content is kept separate from its display. The publisher's job is to describe the content and make it available; the browser's job is to download it and display it. That is the grand bargain that the web offers.

The DOM describes the whole document, and the browser monitors every element in it. The page changes whenever the DOM changes. A WebXR canvas works differently. It is a bitmap, a grid of pixels, and it sits in the DOM as a single canvas element. The browser knows the canvas is there, but it has no idea what the pixels show, so it can't tell a chair from a tree, nor can it combine what is drawn there with content from other parts of the document. This raises security concerns because the site controls every pixel. In immersive mode, a nefarious site could draw a fake login or payment prompt that would be hard to distinguish from a real one.[105]

The SOM gives every owner in a location its own branch of one scene

A webpage has one author, but a metaverse scene has an unlimited number of authors that all collaborate to form a seamless experience. To accomplish this, Sneeze uses a Scene Object Model, or SOM, which is a single tree that the browser understands completely.

Like the DOM, objects in the SOM have a parent and children, but each object's position is set relative to its parent using a three-dimensional transform, so moving a table in a scene also moves everything on top of it. Furthermore, the SOM's job is to hold and arrange all the content in the scene. If you want something to be rendered, you put it in the SOM.

The SOM enforces ownership and privacy

Unlike the DOM, no one location gets exclusive ownership of the SOM. All publishers must share it equitably. Each location can only modify its own branch of the SOM. But any location can see any other location's objects, so long as they're marked as public – which is the default state of an object. Objects can just as easily be marked as private, though. For instance, in a poker game, your hole cards would be marked as private, while everything else is left as public.

While the objects in the SOM are exclusively managed by each location that owns them, the space they occupy is not exclusive. The SOM divides ownership by branch rather than by location, so two

branches can overlap in the scene without either one taking control of the other.

Public content is what lets content from different owners work together. For instance, one owner can publish a warehouse while another provides a way to navigate it. A coffee table can come from one owner while a backgammon board resting on top of it comes from another. That is what makes the metaverse much more than a collection of separate 3D applications.

The renderer reads the SOM, so publishers describe places instead of building engines

The renderer's job is to draw what is in the SOM and make it visible. Sneeze works at the same level as a game engine, so it holds the promise of generating rich, photorealistic 3D graphics. And because the SOM holds all the content in a single, composed scene graph, light and objects from one location can affect any other. For instance, streetlights published by a city shine on the cars and homes around them, which are published by their owners separately.

Publishers no longer have to construct an engine on their own, nor bundle one with the content. With the renderer in the browser, a publisher describes its location and makes it available, the same way they publish a website. The browser on each device fetches it and draws it with that device's own hardware, so a location published once works on a phone, a headset, and a pair of glasses.

Until now, a company had to choose between the reach of the web and the quality of a game engine. Metaverse browsers solve that problem by merging native rendering with the reach and accessibility of the web.



Chapter 15

Spatial Fabrics Are the Websites of the Metaverse

> ***"Someday soon, every place and thing in the real world – every street, lamppost, building, and room – will have its full-size digital twin in the mirrorworld."***
>
> **— Kevin Kelly**
>
> *founding executive editor of Wired*

The website is how the world publishes, and the spatial fabric evolves it for 3D

The website may be the most successful way to publish anything that has ever been invented. Nearly every business, school, hospital, and government agency has one. For simple information, all it requires is an inexpensive domain name, an inexpensive web host, and a few text files – no permission to publish required. That is why the web is so successful.

The web browser makes it possible to do something similar for 3D content using WebXR, but as we've discussed, the limitations of

the browser itself make that an impractical solution for the type of experiences required in the metaverse. It's good for creating standalone custom 3D scenes with only your own content, but monumentally complicated for anything much beyond that.

The metaverse browser, on the other hand, provides the same power as the web browser to self-host and publish spatial content, simply and inexpensively. Instead of building a website, you publish something called a spatial fabric – an openly specified manifest that describes a location, references its content, identifies its publishers and declares the programs that operate on it. It's somewhat comparable to the head of an HTML document.

Spatial fabrics are managed in much the same way you manage your websites. Fabrics are self-hosted on your own server and at your own URL. You can update your fabric whenever you want without going through an approval process like you would in an app store, and without content restrictions like you might encounter when building inside of another company's platform. That also means no company is going to take a large bite of your profits for the privilege of being listed in their store.

Self-hosting meets the enterprise requirements for control and data residency. Open specifications and multiple browser implementations address portability and vendor independence, while link-based and proximity-based access remove the need for a separate app. The company's data stays on its own servers, because the company hosts the fabric itself. Any metaverse browser on any device can open it, so nobody is locked into one vendor's headset or glasses. There are no additional programs to install. Anyone can build tools to make fabric-building easier. Because every fabric follows the same open standards, fabrics from different companies can be composited into a single scene within a metaverse browser. And all this works on the same internet infrastructure that we're already using today.

A fabric represents a location instead of a website

A spatial fabric is a machine-readable manifest for a location. It establishes a coordinate space for a location – such as a planet, corporate campus, or home – or a collection of related content – such

as a game, virtual pet, or a service. It references its 3D assets along with the programs that modify them. And it identifies its publisher using a signed certificate, rather than simply relying on the domain name from which the file was retrieved. The metaverse browser validates that manifest, resolves its resources, builds the permitted elements in the SOM, and renders the resulting spatial content.

The 3D objects referenced within a fabric use popular open file formats that any metaverse browser can read. glTF, the standard format for 3D models, describes what each object looks like. Those 3D assets can be added to a scene via the fabric's running programs or pulled from a metaverse map service – a specialized service designed to communicate specifically with metaverse browsers. Map services are particularly useful for storing very large scenes and often employ a database to manage them. USD is another popular technology that will surely be adapted to store content for map services in the near future.

A fabric can contain sandboxed programs that are executed in the browser, which can modify content in the SOM. These programs often communicate with remote services to provide the data necessary for real-time functionality synchronized for multiple co-present users. There is almost no limit to what a running program can do. For instance, it might act as a visible sensor to report fluid levels on a machine and serve as an instruction guide for refilling the fluid when it's low. Or it could provide information and wayfinding services that operate across an entire corporate campus.

A spatial fabric can be a digital twin tied to an exact location on Earth, but it doesn't always have to represent a real-world location. It could represent something entirely imaginary such as a training floor for a factory that hasn't been built yet or a Zen garden on a planet in a galaxy far, far away from our own. Spatial fabrics work in the browser the same way, the only difference is how you can access them. Real world locations can be experienced with AR glasses when you are present, and other devices such as phones, desktops, and VR headsets when you are not. Virtual world locations can only be accessed with the latter.

If a spatial fabric is comparable to a website, it would be natural to ask what is the metaverse equivalent of a webpage. There really isn't

one. Most websites are divided into pages linked together with hyperlinks.

In the real world, a location isn't divided that way, and neither is a spatial fabric. You navigate continuously from one area into another. Your position changes incrementally, as does the content within the browser that is constantly updating the scene based on your proximity to it. Spatial fabrics are much more like a single-page website that never refreshes, instead continually updating via JavaScript. Anyone who uses a web email client or shared document tool like Google Docs has experienced this.

A location has many owners, so fabrics nest and overlay

A website has one owner, but the metaverse needs to be able to represent real-world locations, and ownership is not nearly as straightforward. Take an office building, for instance. The building itself has a single landlord, but it is filled with tenants who lease office space, which in turn could have sublet offices or leased equipment. Each of them has a need to publish and maintain its own location in the metaverse without handing over control, data, or vital software systems to anyone else.

Metaverse browsers handle this by nesting fabrics via delegation. The owner of the office building publishes the topmost fabric and grants each lessee authority over a defined area or object within it. Thereby, the nested publisher keeps its own fabric, server, and content private. The parent can position the child and apply the permissions agreed in the delegation, but it cannot alter the child's data or view it without authorization. Nesting works a lot like folders on a hard drive.

A delegation identifies the published spatial fabric, the spatial bounds it may occupy, and the capabilities it may request. It can expire or be revoked when a tenancy or commercial relationship ends. A coordinate transform aligns the nested fabric's local map with the containing location.

One advantage of nesting fabrics is that you can compartmentalize reusable fabrics, so a company can create a service that can be used in every branch rather than building that service directly into each one. This is somewhat akin to the iframes that a

webpage uses to show content from another site, but with one important difference. An iframe is a box inside someone else's page, while a nested fabric is a full fabric in its own right, able to do everything that any other fabric around it can do.

Nesting also helps the browser discover increasing levels of spatial detail. Fabrics for buildings, rooms, objects, and services describe what lies within them at incrementally smaller layers. As someone moves around a location, the browser can import only the ones within proximity to them.

An airport is a good example. Many airports are already building digital twins for their facility. The airport runs the terminals, the airlines run their gates and lounges, and dozens of shops and restaurants run their own counters. None of them wants to share their data or run anyone else's systems. The airport publishes the terminal, each airline publishes its own fabric inside it, and each shop publishes its own fabric. A traveler walking to the gate sees all of it as one location, while every owner keeps its content on its own servers.

Overlays are fabrics too. An overlay is content a person enables on top of existing fabrics such as navigation or reviews. They are like adding additional layers on top of the main layer and you can add more than one. Some overlays span the entire coordinate space of another, and others contain personal items that you carry with you from one location to the next. An overlay does not alter the underlying fabric, nor can the underlying fabric alter any of your overlays. But the overlay fabric is a full fabric in its own right, able to do everything that any other fabric around it can do.

Signatures identify publishers across hosting domains

On the web, the authority of a webpage is derived entirely from the origin – scheme (protocol), host (domain/subdomain), and port – that it comes from. A page from one host can't load or share content from another unless the webpage has been preconfigured to allow it, and in some very critical instances they can't share information at all, even when two different hosts come from the same domain. This security model works in a web browser tab that gives exclusive control

to the loaded webpage, but it is not sufficient when dozens of independently hosted fabrics contribute to one scene.

Signatures let the browser associate a fabric with a cryptographic publisher identity even when that publisher uses several servers or hosting providers. They make multi-owner composition safer and less complicated.

Every published fabric is signed with a certificate or verified public key that identifies the signing publisher. While the signature proves that the manifest has not changed since it was signed, it does not prove that the content is safe, truthful, or entitled to occupy a particular location; those questions are handled through delegation, permissions, and policy.

This model can build on existing public-key and code-signing infrastructure, but fabrics still require common rules for certificate validation, expiration, revocation, and key rotation. Reusing established infrastructure reduces what must be invented without eliminating the need for a fabric-specific trust model.

Each fabric or service instance runs inside its own security boundary. The publisher's verified identity informs the browser's trust and filtering decisions, but two fabrics do not automatically share authority merely because the same organization signed them. Any interaction between fabrics occurs only through explicit capabilities granted between them.

Signing isn't required, but an unsigned fabric is treated the way an unsigned website is treated today. The browser can warn you before you go in, but it has no way to tell you who made what you are seeing.

Knowing the publisher lets people and companies decide what reaches them

Fabrics are automatically pulled into the scene as you move around your location without clicking a link. Without a way to filter them, a walk down a busy street could quickly turn into a pop-up nightmare if unchecked. The web learned this late, and pop-up blockers, ad blockers, and cookie banners were each added after the problem had already appeared. A metaverse browser can build filtering in from the start.

Each fabric declares what kind of content it is, such as advertising, social, entertainment, safety, or operations. Webpages already do something similar through Open Graph, a standard that lets a page label itself as an article or a video.[106] A person can switch off a whole category and, for example, never see advertising at all.

A signature does not stop a publisher from mislabeling advertising, but it makes the claim attributable. The browser, the user, or a chosen reputation service can then block that verified publisher, record the violation, or warn other users. Accountability comes from combining signatures with stable identity and enforcement.

Blocking follows the publisher, not a server or an address, so a company can't get around it by moving its content somewhere else. It works the other way too. A person can decide to let in only the publishers they trust, such as large well-known companies like Google, Microsoft, or Meta and keep everything else out until they choose otherwise.

Publishers can set policies for their fabrics. On organization-managed devices, a factory can require approved safety fabrics and suppress advertising. On a visitor's personal device, the factory can request those rules and make compliance a condition of access, but it cannot silently take control of the device. The browser must show which rules come from the user, the location, and the organization operating the device.



Chapter 16

The WordPress Moment for Spatial Computing

> ***"When you build something on it, you own the platform as much as I do… it belongs to you, and no one can take it away from you."***
>
> **— Matt Mullenweg**
> *co-founder of WordPress and CEO of Automattic*

Websites are easy to build because thousands of tools make the same open format

Putting up a website today is almost embarrassingly easy. You can write the code yourself if that is your thing, or you can build an entire site in an afternoon on WordPress, Wix, or Squarespace without writing a single line. Or, if that's too much effort, you could ask your AI agent to do it all for you in mere minutes. You can host it with any one of thousands of companies,[107] add a store, plug in analytics, and hire from an enormous pool of people who already know how to use every one of those tools. WordPress alone powers roughly four in ten

of the world's websites.[108] No one sat down and planned any of this. It simply accumulated, year after year, until building a website became something nearly anyone could do – even Dean's mom. Well, perhaps that's going a little far.

A tool does not need to supply its own browser. It only needs to produce content that conforms to that set of shared standards of a platform. When a tool for the web does so, whatever it produces will work in every web browser on every device. Because the standards are open, nobody needs permission to build a tool like that. So, hundreds of thousands of companies, and millions of individual contributors, did exactly that.

No tool is load bearing, so no tool company can strand you

GeoCities, once home to millions of personal pages, shut down. Flash, which powered much of the animated web for more than a decade, was forced into early retirement. Tools like these for building websites come and go all the time, but other tools – usually better ones – are waiting to replace them. No one tool is the web itself, so the web carries on with or without them.

A closed platform is a different story. There is exactly one set of tools, and it belongs to the company that owns the platform. If that company changes direction, drops a feature you depend on, or raises its price, your work has nowhere else to go. If you need a tool or a feature it doesn't offer, you are at the mercy of the platform – which may or may never get around to building it at all. 8th Wall is a recent example. When it closed its platform on February 28, 2026, developers could no longer log in, edit their projects, or even export them.[109]

Spatial tools today are trapped inside the engine they were built for

Spatial computing already has a handful of powerful creation and publishing tools, but they do not target one widely adopted, vendor-neutral standardized infrastructure like the web. Universal Scene Description (OpenUSD) is the most widely supported file format and is frequently used for data interchange between platforms. It's the

closest thing there is today to spatial infrastructure like the web – but it isn't quite the same.

Blender, a superb free and open-source tool that much of the world uses to create 3D models, is a good example of a 3D authoring tool. It is excellent for making things, but nobody can walk into a Blender model by itself. To become something a person can experience, the objects Blender creates must be imported into an engine such as Unity, Unreal, or Godot where a customized application is built around them. While the assets can be used with any engine, the resulting experience is not transferable. Each engine converts what it imports into its own private format, and the lighting, the behavior, and the services built around the model are custom coded within each engine. A scene assembled in Unreal is fundamentally different from one assembled in Unity or Godot and will not operate with a metaverse browser.

Open formats such as glTF certainly help. They let a model move from one tool to the next, much the way a PNG image moves between photo editors. But a model of an individual 3D object is not a working location. OpenUSD goes considerably further than an asset format. It can compose complex scenes from multiple layers and tools. But it remains principally a production and scene-description system. It does not by itself make independently hosted, positioned services discoverable and executable through every device.

The people who build with these tools are divided the same way. The underlying 3D expertise transfers, but much of the code, project structure and production workflow does not, so the industry's already scarce talent ends up split among engines that don't work together. Web developers similarly have a variety of toolsets to choose from, but their code is portable. A designer who moves between Dreamweaver, WordPress, and Webflow must learn each one, but each produces the same HTML and CSS that every browser reads.

A fabric's programs can be written in almost any language

A fabric doesn't just hold objects. It also holds the programs that manipulate those objects, and those programs run as WebAssembly – better known as WASM. Programmers don't write code in WASM. It's

not a language, but rather a machine interpretable bytecode format their code is compiled into. This is analogous to the way a document can be saved as a PDF regardless of which word processor created it. That means developers can write programs for a fabric in one of several languages they already know. Over three dozen programming languages can produce WASM, including C, C++, C#, Rust, and Go, which are in wide production use, with TypeScript-like AssemblyScript, and many others close behind.[110]

WASM makes the compiled program portable, but the fabric must also standardize the functions that program can call. That interface is the spatial equivalent of the web browser APIs that let JavaScript manipulate a webpage.

JavaScript, the language that brought webpages to life, went down this same road. When it first arrived in 1995, there was almost nothing around it – no libraries, no frameworks, and very few tools. Once it became an open standard in 1997, any browser could implement it and anyone could build for it without asking permission, and developers never stopped. Today its main package registry holds millions of free packages that anyone can use.[111] WASM is open too. It was developed in the open from the beginning and has been a W3C standard since 2019. Companies have been taking advantage of this for years. Adobe and Autodesk both brought major desktop applications to the web by compiling their existing code into WASM. Photoshop and AutoCAD each run in a web browser today because decades of their C++ code was recompiled rather than rewritten from scratch.[112] No doubt, the gamut of tools that exist for WASM today will likely look tiny in a few short years.

Most real locations will be scanned, not built by hand

There are three primary ways to make a spatial fabric: build it by hand, piece by piece, which takes time and skill; scan it; or generate it with AI. For real locations, scanning is quickly becoming the easiest of the three. A person can walk through a room holding a phone with a LiDAR sensor, which measures depth, or fly a drone over a building site, and come away with a 3D copy of it in minutes.

More and more of these scans are captured as Gaussian splats. Rather than building a model out of flat surfaces, a splat scan is made of millions of tiny colored points that, together, look like a photograph you can walk through. Splats capture the kind of detail that older scanning methods struggle with, such as thin wires and shiny metal.

Splats also now have an open standard. In February 2026, the Khronos Group, the organization behind glTF, announced an extension to glTF for storing splats, backed by NVIDIA, Autodesk, Esri, Niantic Spatial, and others.[113] That means a scan made with any app can be placed into any fabric the same way a photo taken with any camera can be posted on any website.

Cesium can already stream very large Gaussian splats through hierarchical 3D Tiles datasets, with nearby details loaded at higher resolution. This capability is also being incorporated into the proposed 3D Tiles 2.0 standard.[114] The part of a location closest to you arrives in full detail, and the rest fills in as you move through it.

Metaverse browsers will only improve at rendering scenes. Because the browser decides how content is displayed, it can improve how a location looks without its owner changing a thing. A scan of a real building published as a spatial fabric and connected to the building's sensors becomes its digital twin.

AI will generate locations the way it already generates video

In the movie Inception, an architect reimagines a city with its streets folded up and over her head. That is not far from where 3D creation is heading. Anyone can already type a sentence or provide a starting photo and have an AI model turn it into a short video clip. World models do the same thing, but in three dimensions. Give one a description or a handful of photos, and it produces a photorealistic scene you can walk through. An AI model can not only build a scene, but it can modify one too. For instance, you could ask it to swap out a sofa for some chairs or redecorate an entire kitchen. It is the one-click builder taken to its logical conclusion.

There is one important distinction – for now. Today's generated scenes are not yet spatial fabrics. Once the AI models are trained to understand the spatial fabric open standard, you'll be able to have the

AI model manipulate spatial fabrics live, which will update live in the metaverse when viewed in a metaverse browser.

Open tools let spatial content grow the way the web did

W3Techs tracks the technology behind the web. They monitor more than a thousand different content management systems, which represent just one category of website tool.[115] The web garnered this level of participation because millions of people, using hundreds of thousands of competing tools, built it together with a shared interest in having an open information-sharing platform.

The spatial fabric gives spatial content a starting point, one open format that every tool can aim at. Add scanning that captures a room in minutes and AI that builds a location from a single sentence, spatial content could easily gain adoption exponentially faster than the web.


91 cm
46 cm

Chapter 17

Spatial Fabrics Should Control Their Positioning

> ***"To find where you are going, you must know where you are."***
>
> **— John Steinbeck**
> *writer and novelist*

Physical content needs a physical position

When you use a desktop or mobile phone, your browser does not need your physical position to render a webpage. A website such as a mapping site may ask for it, but the page itself is not anchored to the room around you – your position in that map is ancillary. The same is true when you visit a wholly virtual location. You pick where you want to go, and the browser takes you there, virtually.

The metaverse and the real world align very differently. A metaverse browser is already able to read the position and orientation of all objects and services in a spatial fabric. It also needs to know exactly where you are within the fabric and which direction you are

facing. Together those two things are called a pose. The same is true for remote visitors joining a real location. If they are not properly oriented within the fabric's coordinate system, they might end up standing inside a wall where nobody in the room will see them.

Poses are not just for people. Robots and machines need to know where they are, down to the centimeter. Within a spatial fabric, the map of objects and services will let them know what is in their immediate surroundings, but they still need to be able to reconcile where they are within the spatial fabric, or else the map is of little use to them.

Web browsers already have a Geolocation API. When a website asks for your location, the browser answers using whatever device capabilities are available – GPS, Wi-Fi, cell towers, or a combination of all three. Each response comes with a measure of how accurate it is. A page can request high accuracy when it needs it, but the browser decides for itself whether to provide it based on the needs of the website. For instance, a weather site is happy knowing which city you are in, whereas a delivery app requires a street address.

A metaverse browser requires much more precise positioning information than a web browser to operate correctly. It needs to know your pose in centimeters, not city blocks. If that weren't difficult enough, the browser may also need to switch from one positioning system to another as you walk, without losing track of where you are.

Public maps get you to the door

There are several technologies available to determine a pose accurately enough to be useful in the metaverse. Knowing roughly where you are in the world is the first half of the solution. For that, we have the global positioning system, or GPS. Consumer grade GPS can identify a building, but it's not accurate enough to pinpoint things around it. In open sky conditions, it's only accurate to about 5 meters.[116]

GPS can identify where on the planet you are, but it can't identify what's around you. For that, we have maps of the world. Lots and lots of maps and of every type imaginable. One such map is OpenStreetMap, a freely distributable open map project started by volunteers in 2004, which now has more than ten million registered

accounts.[117] In 2022, Amazon, Meta, Microsoft and TomTom formed the Overture Maps Foundation to build a shared base map rather than each paying for its own. It has since grown to fifty members.[118] These companies still compete, but they compete on what they build with the map, not on who owns the roads.

These maps are updated constantly, and they have a fairly complete data set that includes streets, addresses, building outlines, and businesses. This is vital information for street navigation and finding destinations. It's also useful for searching and discovering what's around a given location.

What these maps don't have is the detailed geometry needed to navigate the inside of a building and to locate objects and services within it. And GPS lacks the accuracy to be of any further value indoors. The browser, in conjunction with a discovery service, does have the information it needs to retrieve spatial fabrics within proximity. These are handed off to another suite of technologies that provide the second half of the solution for determining an exact pose.

The owner's map takes you inside

A local positioning system such as VPS (Visual Positioning System) takes over where GPS left off. It is used to more precisely determine a pose by comparing what various sensors see or measure against a specialized scan of a location. Several industrial technologies make this possible at meaningful precision including LiDAR, Ultra-wideband (UWB), and acoustic/ultrasonic. For consumer use, the most practical is camera-based VPS.

Each of these solutions has advantages and disadvantages, and no one technology will always work better than the others in all circumstances. The metaverse browser needs to be flexible enough to use any one of them as specified by a given spatial fabric.

Scanning a location can produce two related outputs. One is a positioning map, which includes a set of recognizable features anchored to the location's coordinate system. The other is starting geometry for a digital twin. That geometry becomes a twin once objects are given meaning, connected to the live location, and kept up to date.

When a device enters a location, its camera or other sensors observe the surroundings. The VPS matches those observations against the positioning map. Because the mapped features are tied to known coordinates, the match produces a pose – an estimate of the device's position and orientation along with uncertainty and timestamp values. With a known pose, the metaverse browser can align the device to the coordinate system of the spatial fabric, and the content and services within it.

This process is only as accurate as the current map. Shelves move, machines are replaced, and a construction site can look different every week. If too many mapped features have changed, positioning may become inaccurate or fail altogether. A stale map is particularly dangerous if a robot treats its pose as reliable.

The owner chooses the system and the precision

Different spatial fabrics have different needs for positioning. Some may have privacy requirements to contend with, and others might have operational requirements that demand more precision. Because of this, the browser cannot assume that a single VPS solution will work for everyone. Each spatial fabric has the option to specify one or more VPS solutions to use within its boundaries. If left blank, the browser will simply continue to use the VPS service of its parent's fabric.

Spatial fabrics must have the option to support more than one VPS solution because a variety of devices with different sensors might be navigating the fabric. For instance, a person walking around a factory with smart glasses might only have a camera on their device, giving them less than five-centimeter precision. But a purpose-built robot moving around the same factory might be equipped with ultrasonic sensors giving it sub-centimeter precision. For precision work, like building an aircraft, specialized optical equipment exists that can be accurate to a fraction of a millimeter.

The browser can select the most accurate option available, or let you choose from available options, like the way your browser currently lets you select a microphone or speakers.

Sneeze brings every system together through one API

Most of the technology needed to implement VPS in the metaverse already exists. Google's Geospatial API is built on Street View. Niantic Spatial says its VPS covers more than a million activated locations, while its broader Large Geospatial Model is built from more than thirty billion posed images.[119] Apple, Snap, Immersal and Augmented City all have their own approaches, and Placeframe from the Outernet Foundation is an open-source option.[120] For many public spaces, a cloud-based VPS is exactly the right choice.

To adequately provide a flexible VPS solution, Sneeze will have a Geolocation API that gives all these systems and more one common way to interface with the browser. That way, future technologies that provide quicker or more precise VPS can easily be integrated into the browser.

For this to work, every result must use a common structure, whether it came from Wi-Fi or a laser tracker. It identifies the available position and orientation, the reference frame, when the measurement was taken, its uncertainty, the method and provider that produced it, and the version and age of the map behind it. It also must report any parts of the pose it cannot provide.

That common form is what makes the handoffs invisible. You might start with GPS on the street, pick up the building's Wi-Fi in the lobby, switch to VPS once inside and use ultra-wideband at the machine. Four different systems provide a consistent pose so neither you nor the robot beside you notices the switch.

The browser must transform each result into the same reference frame, compare timestamps and uncertainty, and preserve continuity while systems overlap or disagree since VPS isn't always foolproof. A room may have changed since it was scanned, the light may be poor, or the camera may simply not see enough detail. If a device requires one centimeter precision but the system can only manage one meter, a warning must be given at the very least. Inadequate precision could easily become a safety issue for humans, robots, and machines alike.

The person controls the sensors

The software that implements positioning can run remotely, locally, or through a hybrid of the two. Remote positioning captures images from your camera and transmits them from your device to a separate server, which then computes and returns your position. Local positioning keeps the images on your device and computes your position locally.

There are tradeoffs to be made between these two options. While the latter is preferred for privacy, it means your device will have to download the fabric's maps, store them, and spend more of its time computing your position. This results in reduced battery time and slower processing for other operations your device must perform.

A hybrid approach sends a limited representation of the scene instead of the original images. This is a better option from a privacy standpoint, although the representation still needs some protection because it can retain information about the surroundings.

Ultimately, each individual must decide how much or little privacy they are willing to trade for positioning precision. If the available methods a fabric offers for performing VPS are unacceptable, precise spatial content may simply remain unavailable.

That is the consequence of giving both sides control. The owner decides which methods the spatial fabric supports, the individual decides what their device shares, and the browser enforces it. A camera that you wear on your face should never be allowed to send pictures of your surroundings to anyone without your consent.



Chapter 18

Each Person Should Own Their Identity

> ***"On the Internet, nobody knows you're a dog."***
>
> **— Peter Steiner**
>
> *cartoonist for The New Yorker*

The web made it easy to hide

On the web, anyone can create a fake account in minutes. If an account is ever banned from a site, they can create another account just as quickly. This is how virtually every website operates today and is what allows so much of the worst behavior online to continue. Illegal activity, child abuse material, bullying, and rage posting all rely on it. Behind fake accounts and anonymous networks, the people responsible are hard to find and even harder to stop. They know there are rarely any consequences because nothing connects their accounts to a real person.

Fake accounts also defeat age checks. A child can claim to be an adult, and an adult can pose as a child. Many governments now require

stronger age verification, but on the web that often means handing a copy of your ID or a scan of your face to every site that asks.[121]

Accountability must come with identity

The metaverse not only presents an opportunity to start over, it presents an obligation to do so given the nature of how it operates. Knowing what we've learned from the past four decades being online, we need not repeat the web's mistakes. Identity can and should provide both security and accountability, not only to protect people from having their personal information compromised, but to protect them from being abused by others.

Accountability happens when each account traces back, privately, to one real person, and each person has only one identity. With this policy, identities become difficult or impossible to recreate, which creates consequences for those who break rules. If a person is banned for abuse, they cannot simply make a new account and come back under a fresh name like they can on the web. The ban applies to the person, even though the service that issued it never learns who that person is.

This should not become one universal register of everyone, or a ban that follows a person everywhere. No system can prove that one human has only one identity across every country and issuer without creating a global surveillance system. That is definitely not what we are implying. The practical implementation is much narrower. Where accountability is required, a person can prove that they are a unique participant within that trust system, and consequences remain scoped, proportionate, and open to appeal.

Each company still sets its own rules. It can revoke access and turn away anyone who has broken them. Locations where safety matters can ask for stronger proof than a social space does. For instance, a laboratory, school, or factory might require a safety certification or a background check. Enterprises get the same benefit. A contractor arriving at a client's plant can prove her certification and who she works for, without having a separate account on every client's system. Parents also get similar benefits. They can set a child's rules

once, which accompany the child into every location, rather than depending on every service to apply them.

Done carefully, this makes disposable re-entry much harder without forcing a person to reveal who they are. It may not entirely end abuse, but it will go a long way towards curtailing it. It gives issuers and unmasking authorities considerable power. Therefore limits, audits and a right of appeal matter as much as the credential itself.

Identity, personas, and avatars are three different things

Your identity in the metaverse is who you are as a human being. It lets you authenticate yourself without needing to reveal unnecessary personal details. Your identity wallet can prove individual facts such as being an adult or licensed to do a job, without revealing anything else such as your name.

A persona is how you choose to present your identity in the metaverse. Your identity can have many personas, just like you have in real life – the version of you at work is not the one that plays games with friends. The people and services you meet should not be able to trace a persona back to your identity, nor to one of your other personas.

An avatar is a virtual representation of your persona in the metaverse. Your avatar is placed into a shared space so others can see, hear, and interact with you when you aren't present to them in real life. It may be a photorealistic likeness of yourself for a meeting or a stylized character for a game. Your avatar travels with your persona, so you look consistent wherever you go.

Anonymous is not the same as unidentifiable

Anonymity in the metaverse is still important. People should be able to move through a location, join a forum, or play a game without being personally identifiable. Your persona makes that possible. It allows you to determine which version of yourself you present to the world. People have a variety of very valid reasons for not wanting to be recognized, such as keeping your work and personal hobbies separate or visiting a support group. People should be free to engage in the metaverse without being linked to their real identity – under most circumstances.

Being anonymous to the people around you, however, is different from being unidentifiable to everyone. When you interact with services, there are some legitimate reasons why they might need to know some things about you. For instance, when you purchase a product, the retailer will need to know where to ship it to. Or if you want to gamble, in legal jurisdictions of course, the service will need a way to report annual earnings for tax purposes. If this is implemented properly, the gambling service should be able to do that without knowing anything about you including your name, tax ID, or the country you live in. They don't even need to know your exact age – only that you are legally old enough to play.

The only way to connect a persona to a real person is through the law, with the same kind of due process needed to obtain someone's phone records today. That power must be handled carefully. Whoever can unmask a persona can also abuse that ability, so it needs proper legal process, a right of appeal and penalties that fit what was done. The technology to make this work is not difficult. Deciding who is allowed to use it, and under what rules poses lots of challenges.

Self-owned identity alone makes the problem worse

The most likely candidate for managing identity in the metaverse is the decentralized identifier, or DID. A DID is an identifier that uses cryptographic keys to prove ownership, while its creation, modification, recovery, and deactivation are defined by particular DID methods. A DID alone does not prove that its owner is one unique real person.

While DIDs can improve control and portability, they do not solve accountability. In fact, they make things even worse than they already are today – if you can imagine that. If a system accepts unlimited self-created identifiers as people, each remains disposable, even easier to issue, and therefore much easier to hide behind. The metaverse therefore needs credentials or other proofs above the identifier that can establish uniqueness where it matters without turning every interaction into identification or compromising a person's privacy.

DIDs, therefore, should be viewed as a building block, not a complete solution. What is missing is a layer above that ties each

identity to one real person and prevents people from creating more than one, while keeping that person in control. There are several ways to build that layer, and the approach is not settled yet, but the problem must be solved.

Governments are already building identity wallets

Europe has gone further than anyone. Under the eIDAS 2.0 regulation, every country in the European Union has to offer its citizens a digital identity wallet by the end of 2026. It is free, it is voluntary, and it is meant for everyday life, not just banking and government forms.[122]

The wallet ties an identity to a real, verified person, and then shares only what each situation needs. It can prove that you are over eighteen or that you hold a driver's license without handing over the documents behind them. The regulation also lets people use pseudonyms, so a person can prove information about themselves without revealing who they are.

That is very close to what the metaverse needs – one real person behind each identity, and privacy that protects that person in front of it. It is also something an ordinary person can actually use, which cannot be said of managing a DID.

The web is moving in the same direction. Billions of passkeys now let people sign in without a password,[123] and Chrome and Safari can already ask a digital wallet to prove a single piece of information on a person's behalf.[124] Passkeys still sign you in to each site separately, though, and none of this works at walking pace yet.

Identity must be easy to set up and easy to move

Managing a DID means guarding secret keys, keeping backups and recovering everything if a device is lost or stolen. Realistically speaking, only an infinitesimally small percentage of people on the planet have the skillset or temperament to do that. And for the few who bother to try, many will not do it securely. Someone in their eighties cannot be shut out of the metaverse because the setup is too hard. Identity must be practical for everyone to use – even Dean's mom.

Most people will therefore rely on an identity service provider to be a custodian of their identity, like the way they rely on an email or mobile service provider today. That works if the custodian is never able to own or inspect your identity wallet. A custodian may manage the keys to your wallet, but you must still own them.

Your identity, your personas, and your reputation must be fully portable. You must be able to transfer your identity to another custodian, or take it over yourself, without having to start over. Laws must be passed to guarantee this just like they were for porting your phone number when changing carriers. Losing a phone, a pair of AR glasses, or a headset should not mean losing your identity with them. You sign in on another device and carry on.

Your avatar should not belong to a platform

With co-presence being one of the most compelling features of the metaverse, your avatar will be an extension of yourself. It might be highly individualized to represent your real identity, or it might be a character in a game that you identify as. For whatever reasons you may have, your avatar will be too personal and too important to keep locked inside of another company's service. It must be tied to your own identity wallet so that you can bring it with you wherever you go in the metaverse.

Ready Player Me showed both sides of the problem. It allowed people to create one avatar for games and apps from thousands of developers. After Netflix bought the company in December 2025, its public creator and developer services closed on January 31, 2026. Anyone who had already exported a self-contained glTF or GLB asset could keep using that copy, but applications that relied on the hosted creator, APIs, or delivery service lost them.[125]

Co-presence is more than putting two avatars in the same space. People recognize one another through height, posture, gaze, expression and gesture. If those cues disappear whenever someone interacts with a different service, they effectively take on a new persona.

For avatar portability to function, it needs documented formats, replaceable services, and a migration path before a provider disappears. The avatar format itself needs to carry a portable core that specifies

what a person looks like, how their body moves, how a smile, glance, or pointing gesture is expressed, what they are wearing, and what they are permitted to do with the model. The same meanings must survive even when the devices are different. A high-end headset might show a photorealistic face and full-body motion, while lightweight glasses show a simpler version. Both should still be recognizably the same person.

Open formats such as glTF and VRM already provide much of the foundation. Every conforming metaverse browser will need to agree to accept the same portable core and fall back gracefully when it cannot display every detail. Standardization should not make every world look the same. It should let each world choose its style without forcing people to rebuild themselves at the door.

Robots and AI agents must disclose identity and authority

Robots need layered identities, not just names. A robot crossing a hospital corridor should disclose the device, its current operator, the software controlling it, its authority and any relevant certification. That information can change independently, so their credentials need expiry, revocation, and an audit trail. Similarly, an AI agent should disclose that it is automated before and throughout an interaction, who deployed it, and whether a human is controlling or supervising it. That disclosure should be visible to people and machine-readable to browsers.

When an agent acts for a person or organization, its authority should be explicit, time-limited, revocable, and logged. Accountability then follows the recorded chain of delegation and the applicable law. It cannot simply be assigned to the user for everything the agent does.

People should not have to infer any of this. They should be able to see whether they are dealing with a person, an automated agent, a teleoperated robot, or a mixture. They should also be able to see who is responsible for it and what it is authorized to do.

Identity must be built in from the start

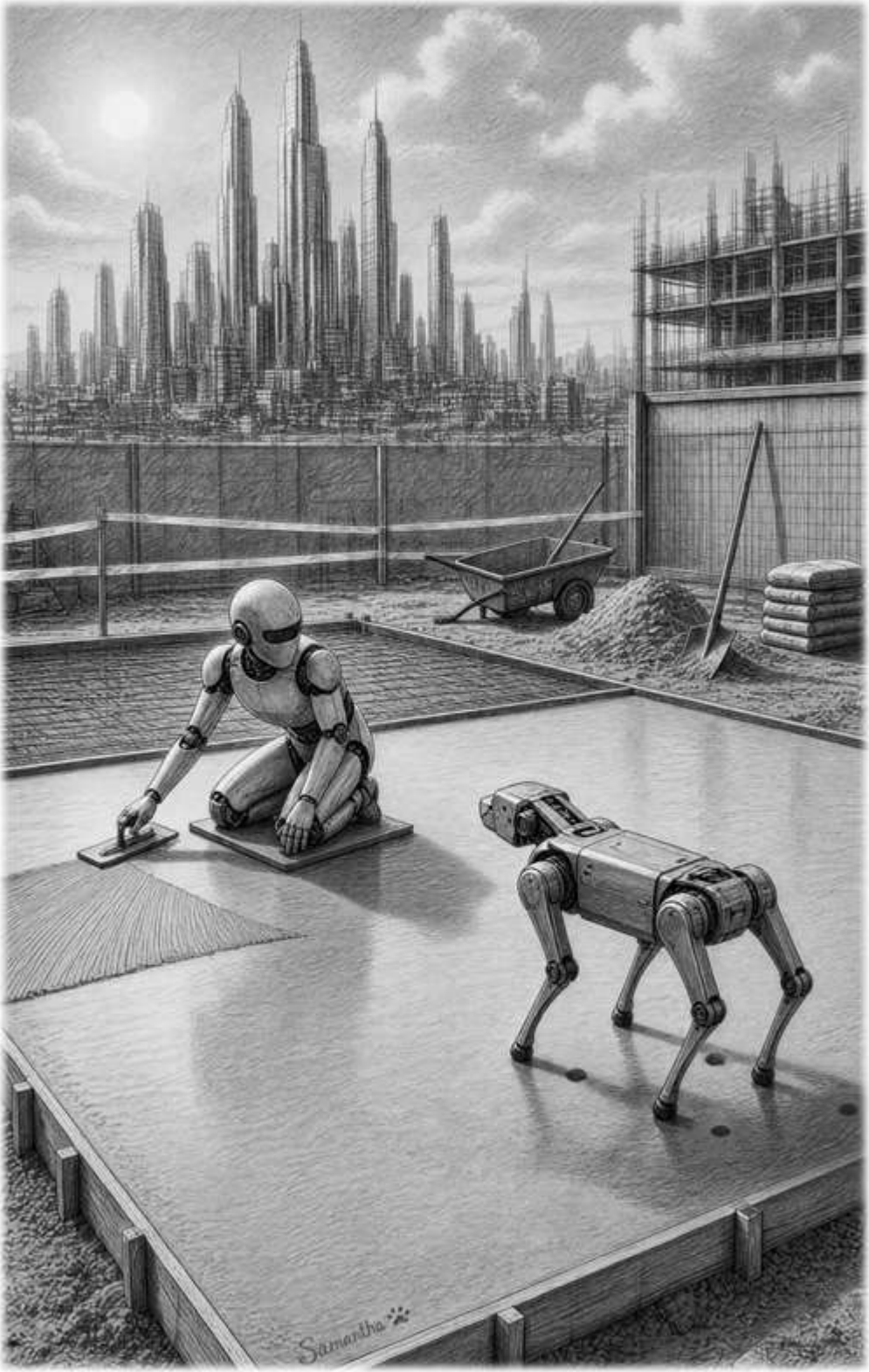
Kim Cameron wrote that the internet was built without a way to know who and what you are connecting to.[126] The web has spent thirty-

five years paying for that with passwords, fake accounts, and abuse it cannot trace.

The metaverse is at its beginning, so identity can be built in from the start rather than added as an afterthought. Done right, a person can be private and accountable at the same time, which the web has never managed. The approach is still being decided, but this must be figured out sooner than later. It won't be long before we're all inhabitants of one, big, open metaverse.

Part 5

Everyone Helps Build the Open Metaverse


Samantha

Chapter 19

RP1 Poured the Foundation for the Metaverse

> ***"Don't worry about people stealing an idea. If it's original, you will have to ram it down their throats."***
>
> **— Howard Aiken**
>
> *designer of the Harvard Mark I computer*

The industry thought a metaverse at global scale was decades away

We started RP1 in 2020, more than a year before Facebook renamed itself Meta and the word *metaverse* became a fixture of quarterly earnings calls. We had technology that allowed a massive number of people to be co-present in an online world – many orders of magnitude more than anyone else – and we wanted to bring that experience to market. Each avatar would be tracked with six degrees of freedom and heard through spatial audio. Our mission was to bring the whole world into the metaverse, not just a dozen or so people in a meeting room.

Article after article assured us the hardware needed to do that was a decade or two away. In December 2021, Intel's head of accelerated computing put a number on it. A metaverse "accessible by billions of humans in real time" would require, he wrote, "a 1,000-times increase in computational efficiency from today's state of the art."[127] Ironically enough, we were already doing that with Intel CPUs that were made in 2014. Two months later, Protocol reported that the metaverse would "require computing tech no one knows how to build." Its conclusion was that the technology might not exist even by 2032.[128] In their defense, we were not very well known at the time.

The Intel post also acknowledged the role of software. Beyond better chips, it said, "we also need new algorithms and software architectures." That sentence was nearly accurate. We didn't *also* need new algorithms and software architectures. We *only* needed new algorithms and software architectures – the hardware has been sufficient for well over a decade.

The cost of scale made walled gardens easier to justify

Those predictions sounded reasonable. Every person in a shared 3D space generates a constant stream of movement and audio. Every other person nearby needs to receive that stream, correctly positioned, at least 60 times per second. As the crowd grows, the work grows exponentially. Using conventional architecture, a space like that is enormously more expensive to host than a website. Very few companies have the resources to run a system as technically challenging as this.

For the companies that can, how many people can they concurrently support in a single location today? Fewer than most would guess. Apple's Vision Pro lets five people share a space as spatial Personas.[129] Meta's Horizon Worlds held 20 people per world when it launched. Meta claims their recently rebuilt architecture can support 100 or more people in one space. But the actual number depends on how optimized the space is. Members-only worlds are still capped at 25 visitors at a time.[130] VRChat has spent years refining exactly this kind of experience, and it caps an instance at 80. Its own engineers have

reported that stress tests began degrading at around 250 people and became "basically unusable" at 400.[131]

The larger numbers that circulate in the industry usually measure something else. Snap's Spectacles allow up to 64 people to share an augmented-reality experience, but only when all of them are standing in the same physical room. Nobody is being transported anywhere.[132] Popular online games routinely put a thousand or more players on one server. However, those are low-fidelity games that don't track each player's head and hands, nor do they provide spatial audio.

For high-fidelity experiences that include spatial audio and six degrees of freedom in one continuous remote environment, the practical ceiling across the entire industry is still measured in the low dozens. Even where games do reach larger numbers, they get there by splitting people up into small groups – also known as sharding. For example, a game of Fortnite can put 100 players into a single game, but those 100 are cut off from the millions of others playing at the same time, each group in its own separate copy of the game and unable to communicate with anyone in any other group.

The real world doesn't work that way. When you walk down a busy street, you can meet and talk to everyone you pass along the way. In other words, the real world is unsharded, and the metaverse must be unsharded too so that anyone in any spatial fabric can meet anyone else, wherever they happen to be, even if they're joining remotely from a VR headset or laptop.

With the best of today's technology measured in dozens of people, or a few hundred when the fidelity is reduced, it's easy to understand why most experts concluded that an unsharded metaverse was decades away. Our technology can change that by making an unsharded metaverse possible.

A different software architecture made large, shared spaces affordable

Dean had been working on this kind of problem long before RP1 existed. Back in 2014, he had built a real-time platform that could support more than a hundred thousand low-fidelity players at once that required only a small handful of CPU cores. The process involved

solving each scalability roadblock, one at a time, like peeling off layers of an onion. At some point he simply stopped because there appeared to be no practical need to improve the system further. There weren't enough players on the planet to fill one machine.

The architecture required between 200- and 500-times fewer servers than conventional approaches that performed the same work. Dean estimates there is opportunity to squeeze an additional 20 times more efficiency from the CPUs on top of that. This architecture permits servers to be synchronized in various ways so that as many as needed can work together as a single unit. Potentially millions of them all coordinating in real-time. None of this relied on exotic processors or ridiculously expensive GPUs. He was simply looking at software differently.

Using far fewer servers meant that a small organization such as ours could afford to host its own part of a shared spatial environment. It no longer had to depend on one platform company to host the whole world. At this early stage in our company, however, we weren't thinking about the web or standards.

The customers we expected did not need that much scale

Our first demonstration ran on a single server that was already several years old – a dual-processor Xeon machine from around 2016. It could support more than 4,000 high-fidelity avatars in one shared space, including spatial audio mixed on the server for every user. Next, we linked servers together and put 100,000 high-fidelity avatars into one continuous shard. It ran on 33 servers, with nothing outlandish under the hood. The only thing that prevented us from going further was our budget.

Most visitors to our early demo arrived skeptical, to say the least. Their skepticism usually turned into excitement with a healthy dose of confusion. Very few people at the time could picture why anyone would need a hundred thousand people in one space. Game developers, the most obvious first customers, had very few games that needed anything close to that scale. Instead, what they really wanted to know was, could they run it on their own servers.

Adding services revealed the limits of the web itself

Two chance meetings – one of them was with Yin-Chien "Yinch" Yeap – led us to choose WebXR for our demo rather than a game engine, having considered both Unity and Unreal. We settled on WebXR because it ran inside an ordinary web browser and it didn't require anyone to download and install any software. Over time, and after countless meetings with a variety of companies, we came to realize that our decision to choose WebXR was just as significant, if not more, than our scaling technology. Few if any companies needed scale, but they all needed solutions that pointed in the direction of the web.

Here we are, sitting on technology that could reduce data center hardware and energy requirements by a factor of 200-500, and no one seemed to care at all. In so many words, what we heard was, *What? You're saying we can save $200 million dollars a month on cloud hosting fees? Nope, not interested in the slightest. But we can't figure out a solution that will allow us to host spatial data on our own servers. What can you do about that?*

We drafted a white paper in 2024 proposing that the metaverse was directly comparable to the web. That 3D spatial fabrics could operate the same way that 2D websites work. That people could publish and host their own spatial data the exact same way they currently host their own informational data. That provided feasible solutions for all the non-negotiables that enterprises have been unable to solve. That, by and large, the metaverse works the same way the web currently works today.

We had to build a new metaverse browser

WebXR, in principle, had it right all along. We've been huge proponents of it since our first decision to use it for our demo. As we tried adding more and more capabilities to it, we discovered that the 3D graphics, admittingly sub-par, were not the thing holding us back. The browser was our primary obstacle.

Things like camera access, pass-through mode, restrictive threading, and limited memory, those were mere annoyances. Some easier than others to handle, but ultimately were not the make-or-break issues in the grand scheme of things. The pathway through the metaverse runs directly through services. Services require multi-origin.

And there was no way in the world the web browser would ever support multi-origin. No matter what we could achieve with WebXR, RP1 could never be more than a demo on a web browser, because web browsers are foundationally architected to prevent metaverse browsing.

By 2025, RP1's visibility was growing and our message was beginning to take root and sprout amongst executives at both the technology producing companies and the technology consuming enterprises. Employees at many of the largest companies and universities were regularly telling us that not only does this make perfect sense and solve all the problems they've been blocked on for years, but it's the only way it could ever possibly work. Once you see the vision, it's impossible to unsee.

We started exploring ways to open-source pieces of our technology, knowing that we were getting ready to embark on building a metaverse browser and engine. At that time, one of our biggest proponents was Neil Trevett of NVIDIA, The Khronos Group, and the Metaverse Standards Forum, who convinced us that the MSF was the right home for that project. The MSF brings standards organizations together, which made it a natural fit for something this ambitious. For a team our size, it was an extremely exciting step to open up the work and invite the rest of the industry to help build it, and a chance to help unite an industry that had spent years building in different directions.

The Open Metaverse Browser Initiative was officially formed under the MSF in March 2026. And coding on Sneeze and Rubidium began in earnest at the same time. Within three months, Sneeze was officially announced in June at AWE and Rubidium a few months later. As if our small seven-person team didn't have enough to do, here we are in October publishing a book to explain the monumental importance of these two products.

The browser needs support beyond RP1

Sneeze and Rubidium are not meant to be a better demo. They are meant to be the foundational engine and browser that the metaverse will eventually be built upon. At the time of this book's publishing, they

can browse the metaverse. Sneeze has a solid multi-origin core that supports metaverse browsing of multiple fabrics based on proximity, as described herein. You can build networked services, manipulate the SOM, and of course, load content that you publish on your own servers.

Believe it or not, inventing and building technology that no one else thought was possible for two more decades was not the difficult part. We had that done within a few months of founding our company. The hard part was first figuring out for ourselves how it was supposed to be used, and then convincing the rest of the world that we shouldn't abandon the web in our pursuit to build the metaverse. Explaining that the web is the best path for the future has been surprisingly difficult. Unimaginably difficult.

This is not the end of our journey by any means. We just can't go much further alone. At times, being a small company was an asset. Three programmers can make quick decisions, and we had no existing platform or business model to protect. But building a browser is not a job for three programmers. Three thousand seems more appropriate. Full-fledged, industrial scale browsers require more than basic operability. And quite frankly, Dean really needs a vacation.

RP1 has poured the foundation of the metaverse. We have donated years of our ideas and code from several open-source projects that will eventually support the full weight of the metaverse. Other companies and institutions now must help develop and standardize the technology for it to reach its full potential.


Sass

Chapter 20

Together, We Can Build the Metaverse

> ***"The Internet is for everyone, but it won't be unless we make it so."***
>
> **— Vint Cerf**
> *co-inventor of the Internet*

Open foundations grow through shared action

The web did not become the foundation of the world's information because its standards were elegant. It became the foundation because billions of people used it and made things with it – webpages, businesses, tools, and communities – until it was too ubiquitous to ignore.

Here are a few of the companies that were some of the earliest to recognize the power of the web. Amazon opened in 1995, when most businesses had never heard of a website. Google arrived in 1998, years after the first search engines, and Salesforce began selling business software through a browser in 1999, when serious business software

still came on a disc. WordPress followed in 2003 and Amazon Web Services in 2006. None of these companies was first, but they were all innovators. And by being early, they had less competition, so their innovation propelled them to the top where they all still stand today.

The web has an unusual way of bringing fierce competitors together to form alliances that allow them to prosper in ways they could not achieve on their own. Apple, Google, Microsoft, and Mozilla form the WHATWG Steering Group, while Interop brings those companies and other browser-engine contributors together around shared features and tests intended to improve cross-browser compatibility.[133]

Working together on open standards such as the web browser might seem altruistic, but like all business decisions, companies do it because it's much better for their bottom line. Rivals agree on the foundation they stand on and then compete over everything built on top of it. This is advantageous because the market created by an open web is much larger and more durable than any closed platform they could own outright.

The metaverse hasn't quite developed its foundation – yet. That's potentially because most people haven't read this book – yet. In the interim, companies have resorted to building their own platforms instead of building their businesses on top of the foundation of the metaverse. The companies that help lay that foundation will be the ones standing atop it when the market arrives.

The foundation of the metaverse will be laid by three forces in particular – the technology companies that build products and services, the enterprises that buy them, and the public institutions that help set and test the rules, including standards bodies, governments and universities.

Browser companies can bring metaverse browsing to everyone

If you've ever been part of a team that builds web browsers, you understand the importance of the browser engine. Sneeze is designed to sit right beside the engine your browser already ships, so webpages will continue to load exactly as they do today and spatial fabrics will use the new engine. WebXR remains intact and those who've helped build it will recognize Sneeze as a torchbearer of the principles behind it.

Google and Meta are in a particularly strong position to lead, because each of them already builds both a web browser and the glasses people wear that will need a metaverse browser. Either one could make metaverse browsing an ordinary part of smart glasses within a single product cycle, reaching everyone who already uses its browser on the phones, laptops, and headsets they own today, long before most of them buy a pair of glasses. The companies that ship metaverse browsers first will play a much bigger role in deciding how they work.

The W3C has a natural part to play too. It is where browser companies already meet to agree on how the web behaves, and it has already heard the case for a metaverse browser engine from the Open Metaverse Browser Initiative. Other layers will involve WHATWG, Khronos, IETF, OGC and other standards bodies. Recognizing metaverse browsing as part of where browsing is headed and chartering work on how web browsers and metaverse engines should cooperate would signal to the world that this is the foundation to start developing for right now.

Smart-glasses makers can compete on the device

A device is only as valuable as what people can do with it. Right now, every company making smart glasses is expected to supply that value by itself before anyone has a reason to buy the hardware. This means building an entire ecosystem, complete with its own store, its own platform, and its own catalog of experiences. Very few companies have the capital to keep paying for a closed software ecosystem until it pays for itself, and most smart-glasses makers are not among them.

The personal computer industry faced the same problem once. In 1987, IBM introduced a new proprietary design for the inside of its PCs and began asking the companies that built compatible machines to pay licensing fees to use it. The following year, Compaq gathered eight other manufacturers, from Hewlett-Packard to Tandy, and together they announced an open alternative that any company could build on for a nominal fee. More than sixty companies signed on.

Individually, none of them could match IBM. Combined, they were selling more personal computers than IBM was, and over the following years, PC standards were transferred from a single company

to the entire industry to manage together.[134] In 1983, IBM owned 63 percent of the entire PC market. In 1996 it was under 10 percent.

Smart-glasses makers have the same opportunity today. With Sneeze embedded, your glasses can reach any location, service, or digital twin that hundreds of millions of people will one day publish. You can compete on your core competency – hardware, optics, and the browsing experience.

The five requirements enterprises have set for spatial computing also hand you a specification to build toward, because glasses that work across vendors and let their owners keep their own data are glasses a hospital or a factory can standardize on. The most useful step now is to embed the engine in a prototype, test metaverse browsing across other companies' devices, and help define what the engine needs from the hardware underneath it.

Toolmakers and service providers can build the markets above the standards

No ecosystem forms without tools that publish to it, and the most valuable feature an authoring company can add right now is a way to publish a spatial fabric. A scene built in one engine's editor stays in that engine, while a fabric exported once can be opened by any metaverse browser and on any device. Because a fabric is self-hosted like a website, the people who use your tools own what they make and are less exposed to a platform closure taking their assets with it.

Almost every service the web depends on has a spatial counterpart. Hosting companies can host spatial fabrics for organizations that would rather not run their own servers. Someone also needs to build better map servers that store and serve those spatial fabric maps at scale. OpenUSD, a widely supported format for composing and exchanging complex 3D scenes, is a natural foundation for a metaverse map server. Search companies can build the index that lets people and AI agents find what isn't in front of them, identity and trust providers can issue and verify the credentials browsers rely on, and integrators can connect fabrics to the systems enterprises already run. Many of these roles still lack a clear leader.

Implementers and standards bodies can turn working code into standards

It is unlikely that the standards the open metaverse requires will be written in advance and then implemented. They will be discovered by companies building working products with the open metaverse browser engine and testing them against one another. The Open Metaverse Browser Initiative is where those findings are coordinated, and the standards organizations (whose job this actually is) then decide which existing standards to extend and where new ones are needed.

The most concrete contribution a technology company can make right now is to nominate a few engineers, build one reference integration between your product and someone else's on Sneeze, and publish the result so the standards bodies have something real to work from. If you are a chipmaker, a cloud provider, an AI lab, or a robotics company, you are probably spending billions of dollars on the future this layer connects. Supporting this work is the cheapest insurance any of you will ever buy on the money you are already spending. Standards are not built by spectators. They are built by those who show up first.

Enterprises can make openness a requirement

Enterprises are the ones who taught this industry the word *no*, and you were right to do so. Platforms you invested in were discontinued, deployments were stranded, and you learned not to put critical operations on infrastructure you do not control. The open foundation you have been waiting for will soon be ready to use. You can already start developing on it right now.

Your buying power is one of the strongest forces that can pull this industry toward open infrastructure. Use it. Write openness, self-hosting, and interoperability into the agreements for the systems and services you buy, and ask every supplier, "What happens to our deployment if your company shuts it down?"

Go one step further than procurement. Publish a spatial fabric of your office or corporate campus on your own servers, the same way you once put up your first website. Assign a few of your technical staff to the working groups, because nobody knows better than you what the standards need to say about security, compliance, and day-to-day

operations. Hospitals, airports, and utilities buy infrastructure meant to last for decades. The market will build for what they require.

Governments can make open standards public policy

History suggests that we can't rely on an industry to police itself. Roads, the electrical grid, radio spectrum, and the internet itself were all built with governments having a seat at the table. Decisions regarding critical infrastructure the public depends on cannot be left to whichever company happens to own it. Governments worldwide will need to be involved in the standards that bring the physical world online.

There are three specific ways governments can help. The first is to procure systems based on open standards. In the United States, federal agencies are already directed by law to use voluntary consensus standards in their procurement rather than inventing their own.[135] Applying that same principle to spatial computing would make every public hospital, transit system, and military base a buyer of open infrastructure. The second is to show up. National standards agencies can join the working groups where these standards are being defined, so that public safety, privacy, and accessibility are designed in from the beginning rather than regulated in afterward. The third is to fund the standards bodies and research that larger companies don't write checks for and smaller companies can't afford. Much of this research happens in universities.

Universities can advance what is still unsolved

There is no shortage of questions and problems that remain unsolved for the metaverse that span graphics, networking, security, AI, human-computer interaction, and public policy. Because the foundation is open, where academia works best, opportunities abound to publish, build reference implementations, train the people who will carry this forward, and shape the standards while they are still soft enough to shape.

This endeavor at the university level is already underway. The University of Rochester, through its Center for Extended Reality, started the Open Metaverse Academic Alliance (OMAA) to bring

universities together around this open foundation. More than forty institutions have already joined.[136] The web itself was not born in a company. It grew out of the research community, and its early standards were worked out and legitimized by universities that had no incentive other than to get them right. The open metaverse is drawing the same kind of academic gravity for the same reason, and joining the alliance is one of the easiest ways for an institution to make an impact.

After you join OMAA, publish a spatial fabric of your campus on your own servers, just as universities ran many of the web's first websites. Then send a few graduate students to the working groups, so their research feeds directly into the standards being written.

Every participant can start with one working piece

You do not need permission or a complete metaverse strategy to begin. Every one of these roles can start with something small enough to finish in a single quarter – one prototype, one building, one service, one tool, or one integration. The fastest way to lay the foundation is to build a working piece of the metaverse. The code, the documentation, the browser, and the ways to join the work can all be found at omb.wiki. Once you've built something and experienced it firsthand, the next most valuable thing you can do is convince someone else.

The web grew because anyone could build on it

We began with a misunderstanding, the idea that the metaverse was a virtual world you escape into. But it was always something larger and more tangible than that. It is the web, evolved into the physical world, connecting our locations, our things, our machines, and our presence in the same way it once connected our documents.

Standards work is slow and easy to skip. It is also how the world has built, time and again, infrastructure that companies and countries can rely on without depending on any single owner. We now get to make that choice again, deliberately this time, with history as our guide.

The web did not grow because one company built it well. It grew because anyone could build on it. The open metaverse will grow the same way, one browser, one device, one building, and one tool at a

time. The people who will have the most impact will be the ones who decided not to wait for someone else to start. Please be one of them.

Metaverse
STANDARDS FORUM™

Join the Metaverse Standards Forum

OPEN METAVERSE BROWSER INITIATIVE

The open metaverse needs your voice. Join us.

JOIN OMBI

University of Rochester

Open Metaverse Academic Alliance

Bringing Together Academic Institutions Worldwide Around New Open Standards for the Metaverse Browser

Are you with a University? Join Here.

JOIN OMAA



Endnotes

All references used in this book are listed on our website with additional detail and hyperlinks to source material for easy access. Scan this QR code or type the URL below in your web browser.



https://rp1.com/book/the-spatial-age

[1] Neal Stephenson, *Snow Crash* (Bantam Books, 1992).

[2] RAND Corporation, "The Metaverse: What It Is and Is Not," commentary, June 2022.

[3] Frank Rose, "Keyword: Context," *Wired*, December 1996. Widely repeated secondary accounts date the line to 1994.

[4] Matthew Ball, "Framework for the Metaverse," June 2021, expanded in *The Metaverse: And How It Will Revolutionize Everything* (Liveright, 2022). Ball's definition continues past the phrase quoted here to require synchronous persistence, an individual sense of presence, and continuity of identity, history, entitlements, objects, communications, and payments.

[5] Meta, "The Facebook Company Is Now Meta," 28 October 2021.

[6] Meta Platforms, *Form 10-K, Reality Labs segment results*, fiscal years 2020 through 2025. The cumulative figure is the sum of the annual segment operating losses as reported.

[7] Siemens, "Introducing Digital Twin Composer," and Siemens partnership materials with NVIDIA.

[8] Statista's global datasphere series, drawing on IDC's Global DataSphere, puts data created worldwide in 2025 at roughly 181 zettabytes. IDC's measure covers data created, captured, copied, and consumed.

[9] BMW Group, "Innovative '3D Human Simulation': BMW Group Plant Regensburg Uses Virtual Tools to Plan Assembly Processes Years Ahead of NEUE KLASSE Series Launch."

[10] Netcraft, "July 2026 Web Server Survey," 27 July 2026, and Verisign, *Domain Name Industry Brief*, second quarter 2026. Netcraft prints no active-site total, and the figure is derived from the percentages in its active-sites table.

[11] Xiao Xiang Zhu et al., "GlobalBuildingAtlas: An Open Global and Complete Dataset of Building Polygons, Heights and LoD1 3D Models," *Earth System Science Data* 17 (2025): 6647-68, doi:10.5194/essd-17-6647-2025, and Laurens J. N. Oostwegel et al., "From Footprints to Functions: A Comprehensive Global and Semantic Building Footprint Dataset," *Scientific Data* 12 (2025): 1699, doi:10.1038/s41597-025-06132-z. Both draw on Google Open Buildings, Microsoft Global ML Building Footprints, and OpenStreetMap, so the two counts are not wholly independent.

[12] Sandvine, *The Global Internet Phenomena Report*, January 2023.

[13] Reuters, "AI Pioneer Fei-Fei Li's World Labs Raises $1 Billion in Funding," 18 February 2026; SiliconANGLE.

[14] World Labs, "Bigger, Better Worlds."

[15] World Labs, "World Labs Acquires SceniX," 21 July 2026, and "Building Worlds That Train Robots."

[16] AMD, "AMD to Acquire World Labs to Advance the Future of AI," 28 September 2026, and Advanced Micro Devices, Form 8-K, 26 September 2026. The transaction is expected to close by the end of 2026, subject to regulatory approval.

[17] NVIDIA, Cosmos product documentation; NVIDIA Developer Forums, July 2026, reporting more than ten million downloads.

[18] Google DeepMind, "Genie 3: A New Frontier for World Models."

[19] Runway, "Introducing Runway GWM-1"; TechCrunch, 11 December 2025.

[20] Reuters, 19 November 2025.

[21] Crunchbase News, "Robotics Startup Skild AI Triples Valuation," 2026.

[22] Dealroom figures reported by Sifted; New Market Pitch, "Humanoid Robotics Funding Trends," 2026.

[23] Goldman Sachs Research, "The Global Market for Robots Could Reach $38 Billion by 2035."

[24] The Motley Fool, "Tesla's Optimus AI Bet," 11 September 2025; Bank of America sum-of-the-parts analysis placing Optimus at roughly 19 percent of Tesla's value, reported by Teslarati.

[25] Reuters, "Robotics Startup Figure Valued at $39 Billion in Latest Funding Round," 16 September 2025.

[26] "Skild AI Raises $1.4B, Now Valued Over $14B," Business Wire, 14 January 2026.

[27] TechCrunch, "Physical Intelligence Is Reportedly in Talks to Raise $1 Billion Again," and Bloomberg, 27 March 2026.

[28] 1X Technologies, Neo product announcement; Digital Trends.

[29] Apptronik partnership announcements for Mercedes-Benz and GXO.

[30] Quartz, 18 August 2026; CNBC.

[31] Waymo, "Waymo Raises $16 Billion Investment Round," 2 February 2026.

[32] Goldman Sachs Research, "Robotaxis Are Forecast to Become a $400 Billion Market in 2035," which puts the global market at about $415 billion.

[33] Counterpoint Research, on shipments in the first half of 2025; CNBC, 11 February 2026, on EssilorLuxottica's sales. The volume figure is company-reported and covers the Ray-Ban and Oakley Meta lines together.

[34] Meta, September 2025.

[35] Meta, "Introducing Meta VR Glasses: A Cinema, Courtside Seat, and Workspace in Just 100 Grams," 23 September 2026.

[36] Snap Inc., "Introducing SPECS Augmented Reality Glasses," 16 June 2026. The 47 mm model weighs 132 grams and the 52 mm model 136 grams.

[37] Snap Inc., "Snap Inc. Debuts Specs Augmented Reality Glasses to Make Computing More Human," 2026.

[38] Google, "Android XR at I/O 2026"; Samsung Newsroom.

[39] Bloomberg reporting by Mark Gurman, summarized at 9to5Mac, 29 August 2026.

[40] Qualcomm, "Qualcomm Takes Spatial Computing into the AI Era with Snapdragon Reality Elite," Augmented World Expo, June 2026, in which Ziad Asghar said there are "more than 60 million devices already in market"; Qualcomm, fiscal third quarter 2026 earnings presentation, 29 July 2026, on roughly 600 million eyewear units shipped globally each year, with smart glasses at about 1 percent; Cristiano Amon, interview on CNBC's "The Tech Download," 16 June 2026.

[41] NVIDIA, Cosmos and Omniverse product documentation.

[42] GSMA, *State of Mobile Internet Connectivity 2025*, which puts smartphone ownership at 4.4 billion people, or 54 percent of the world.

[43] *Late Show with David Letterman*, 27 November 1995, recounted by CNBC.

[44] Global Business Travel Association, *Business Travel Index Annual Global Report and Forecast*, eighteenth edition, 2026.

[45] Hannah Ritchie, "Global Aviation Emissions," Our World in Data, for the emissions share and for how few people fly. Corporate reduction targets are validated by the Science Based Targets initiative and tracked by the Travel Smart Campaign.

[46] "On the Relative Importance of Visual and Spatial Audio Rendering on VR Immersion," *Frontiers in Signal Processing* 2 (2022), doi:10.3389/frsip.2022.904866.

[47] Nicholas Bloom, Ruobing Han, and James Liang, "Hybrid Working from Home Improves Retention Without Damaging Performance," *Nature* 630 (2024): 920-25, doi:10.1038/s41586-024-07500-2. A six-month randomized controlled trial of 1,612 graduate employees at a single company.

[48] *San Francisco Chronicle* and Bisnow, 28 August 2020, on Pinterest's $89.5 million termination of its lease at the unbuilt 88 Bluxome development.

[49] Jeremy N. Bailenson, "Nonverbal Overload: A Theoretical Argument for the Causes of Zoom Fatigue," *Technology, Mind, and Behavior* 2, no. 1 (2021), doi:10.1037/tmb0000030, tested empirically by Queiroz, Fauville, and colleagues at the Stanford Virtual Human Interaction Lab across 7,846 respondents.

[50] Hannah Ritchie, "Global Aviation Emissions," Our World in Data, for the emissions share and for how few people fly. Corporate reduction targets are validated by the Science Based Targets initiative and tracked by the Travel Smart Campaign.

[51] Apple Inc., Form 10-K for the fiscal year ended 27 September 2025: iPhone net sales of $209.586 billion against total net sales of $416.161 billion.

[52] 42matters counted 2,576,072 apps on Google Play and 2,633,628 on the Apple App Store as of 17 September 2026. Trackers differ in how they count a live listing, and more conservative counts put the combined total nearer four million.

[53] Sensor Tower, *State of Mobile 2026*. That the daily set is largely stable is shown separately by Tu et al., "Real World Longitudinal iOS App Usage Study at Scale," arXiv:1912.12526.

[54] Android Developers Blog, discontinuation notice, October 2018. Google stopped serving the notifications on 6 December 2018, citing "a significant increase in locally irrelevant and spammy notifications."

[55] Airport figures from ATL, "About ATL"; passenger count from Airports Council International World, world's busiest airports 2025.

[56] Apple App Store Small Business Program; Google Play service fee schedule. Google's structure is not a flat thirty percent: it charges fifteen

percent on the first million dollars a developer earns each year and thirty percent above it.

[57] Userpilot and Panto, 2026 compilations of mobile app retention benchmarks. Figures of this kind are drawn from analytics vendors' own panels rather than from a census.

[58] World Economic Forum, "Driving Global Growth Through the Power of Small Business."

[59] IDC, "Smart Glasses Surge," for the shipment figures and the full-year forecast; Gartner, press release of 26 August 2025, for the enterprise agent forecast.

[60] IDC, *Worldwide Humanoid Robotics Market Analysis 2026*, putting total shipments near 18,000 and finding that most went to entertainment, research, education, and data collection. Omdia's count of general-purpose embodied robots is lower, at 13,318, so the figure in the text holds on either.

[61] AgiBot's deployment of nearly 100 A2-W robots at Fulin Precision's auto-parts plant, 2025; UBTech, Walker S1 deployment at Zeekr; GXO and Agility Robotics, "GXO Signs Industry-First Multi-Year Agreement with Agility Robotics," 27 June 2024.

[62] Boston Dynamics, Spot product page, and Swiss National Science Foundation, "Robot Dog ANYmal Gains Momentum."

[63] Boston Dynamics case studies covering Michelin's Lexington facility, National Grid's Sandy Pond converter station, and BP's Mad Dog offshore rig.

[64] Jihan Yang et al., "Thinking in Space: How Multimodal Large Language Models See, Remember, and Recall Spaces," CVPR 2025, arXiv:2412.14171. On the VSI-Bench benchmark humans score 79.2 percent against 48.8 percent for the strongest model tested, and roughly 71 percent of that model's errors are failures of spatial reasoning rather than of visual perception.

[65] Javier Civera, Andrew J. Davison, and J. M. M. Montiel, "Dimensionless Monocular SLAM," IbPRIA 2007. A moving calibrated camera recovers scene geometry only up to a scale factor, so absolute scale requires at least one metric fact from outside the image.

[66] Amazon, "Amazon Deploys Over 1 Million Robots and Launches New AI Foundation Model."

[67] Waymo, "The Waymo Driver Handbook: How Our Highly-Detailed Maps Help Unlock New Locations for Autonomous Driving," September 2020.

[68] Applied Intuition and Digital Divide Data, on the cost, update cycles, and per-vendor duplication of high-definition map production.

[69] Applied Intuition and Digital Divide Data, on the cost, update cycles, and per-vendor duplication of high-definition map production.

[70] ISO 10218-1:2025 and ISO 10218-2:2025, which direct compliance with IEC 61508-2 clause 7.4.11 wherever data implements a safety function. The American equivalent is ANSI/A3 R15.06-2025.

[71] Supply Chain Research, buyer's guide to warehouse robotics, and comparable industry guides, which converge on three to six months.

[72] Safari does not implement the WebXR Device API on iOS, iPadOS, or macOS, and Firefox does not implement it at all. Safari on visionOS enables immersive virtual-reality sessions but not the Augmented Reality Module. WebXR is among the Interop 2026 focus areas, so these gaps may narrow. Meta Horizon developer documentation; Apple Developer Forums; PlayCanvas platform support.

[73] Sergio Villar of Igalia, "WebXR in WPE WebKit," on the cross-process handoffs and explicit GPU synchronization an immersive session pays on every frame; and *Demystifying Web-based Mobile Extended Reality Accelerated by WebAssembly*, doi:10.1145/3618257.3624833.

[74] W3C Geolocation API, Recommendation of 28 October 2025. Accuracy is defined as a 95-percent-confidence radius in meters, altitude may return null, and heading reports the direction the device is moving rather than the direction it is pointing.

[75] Metaverse Standards Forum, "The Case for a Metaverse Browser Engine," presented to the W3C Immersive Web group, New York, May 2026. Its two tables of foundational limitations are the source for this chapter's claims about frame timing, memory and compute budgets, augmented-reality fragmentation, service discovery, addressable live services, and multi-origin composition.

[76] Neil Trevett, in the Open Metaverse Browser Initiative session at SIGGRAPH 2026, Metaverse Standards Forum, 27 July 2026. This is what the initiative's members report hearing from enterprises rather than a measured survey finding.

[77] Unity and Epic Games build-target and packaging documentation.

[78] The two services ran incompatible proprietary mail systems and could reach each other only through internet gateways, which AOL opened in June 1992 and which charged extra, truncated long messages, and dropped

characters they could not carry. AOL press release, 3 June 1992; contemporaneous gateway guides in *TidBITS*.

[79] CNN Money, "AOL and Time Warner to Merge," 10 January 2000, for the market capitalization.

[80] Market capitalizations as of September 2026: Alphabet roughly $4.1 trillion, Amazon roughly $2.8 trillion, Meta roughly $1.6 trillion.

[81] Meta Platforms, Form 10-K, Reality Labs segment results, fiscal years 2020 through 2025. The cumulative figure is the sum of the annual segment operating losses as reported.

[82] Road to VR, on Meta's "over $2 billion" figure for Quest store content sales since 2019, first given in September 2023 and repeated unchanged at the Game Developers Conference in March 2025.

[83] Sony Group Corporation, consolidated financial results for fiscal 2025, Game and Network Services segment, which reported sales of roughly $31 billion for the PlayStation business.

[84] Road to VR, "Meta CTO: VR Gaming 'Gravy Train' Has Stopped, Customer Acquisition Now the Real Problem," 2026. Bosworth said the "gravy train has come to a stop"; the assessment that Meta's funded titles bought goodwill with enthusiasts but did not materially expand the market is Road to VR's.

[85] The Verge, February 2025; Microsoft Learn release notes confirming the device was no longer manufactured as of December 2024.

[86] Road to VR; Niantic Spatial, "The Next Chapter."

[87] Microsoft shut the service down in March 2023; UploadVR.

[88] TechCrunch, 19 December 2025.

[89] Microsoft, with Intel and IoT Analytics, *IoT Signals: Manufacturing Spotlight*, a survey of 500 manufacturers worldwide conducted in April and May 2022, in which seventy-five percent named difficulty of integration among their top challenges.

[90] Microsoft, with Intel and IoT Analytics, *IoT Signals: Manufacturing Spotlight*, a survey of 500 manufacturers worldwide conducted in April and May 2022, in which seventy-five percent named difficulty of integration among their top challenges.

[91] HIPAA Privacy Rule, 45 C.F.R. §§ 164.502(e) and 164.504(e), on business associate agreements; FERPA, 34 C.F.R. § 99.31(a)(1)(i)(B), on the direct-control condition; Gramm-Leach-Bliley Safeguards Rule, 16 C.F.R. § 314.4(f), and the federal banking agencies' "Interagency Guidance on Third-Party Relationships," June 2023; FedRAMP Authorization Act, in the

National Defense Authorization Act for Fiscal Year 2023, and DFARS 252.204-7012.

[92] Simon Kemp, *Digital 2025: Global Overview Report*, DataReportal, with We Are Social and Meltwater, February 2025. The figure comes from GWI's survey of internet users aged 16 to 64 and covers time online on any device, not browser time alone.

[93] *Cambridge Business English Dictionary*, "web browser," Cambridge University Press.

[94] Open Hub, "The Chromium (Google Chrome) Open Source Project: Languages," March 2026. Open Hub counts 49,251,381 lines in Chromium, the open-source project behind Chrome, of which 37,652,156 are code and the rest are comments and blank lines.

[95] The Chromium Projects, "Blink," and the WebKit project.

[96] Meta, "Browser Introduction" and "Browser Specifications," Meta Horizon OS developer documentation.

[97] Metaversal Corporation, Rubidium source repository, GitHub.

[98] Metaverse Standards Forum, "Introducing Sneeze, the World's First Open Metaverse Browser Engine," June 2026. The source code is published on GitHub under the Apache License 2.0.

[99] Tim Bajarin, "At AWE 2026, Spatial Computing Grows Up And AI Is Now The Story," *Forbes*, 18 June 2026.

[100] Metaversal Corporation, Rubidium source repository, GitHub, released under the Apache License 2.0 in September 2026.

[101] Metaverse Standards Forum, "Metaverse Standards Forum and RP1 Create the Open Metaverse Browser Initiative," 11 March 2026. Neil Trevett described the Initiative's structure at SIGGRAPH 2026.

[102] Unity Software Inc., fourth-quarter and full-year 2023 letter to shareholders, February 2024, filed with the Securities and Exchange Commission.

[103] W3C, *WebXR Device API*. A request for an immersive session is rejected while another immersive session is active.

[104] W3C, *Document Object Model (DOM) Level 1 Specification, Version 1.0*, W3C Recommendation, 1 October 1998.

[105] W3C Immersive Web Working Group, "WebXR Device API Privacy and Security Explainer," on trusted UI, and W3C, *WebXR Device API*, "Trusted Environment."

[106] The Open Graph protocol.

[107] HostList, "Web Hosting Statistics 2026," which tracks 7,281 dedicated web hosting companies in 87 countries, and HostDir, "The State of Web Hosting 2026," which counts 4,191 active providers in 170 countries.

[108] W3Techs, "Usage Statistics and Market Share of Content Management Systems," October 2026. W3Techs finds WordPress on 40.2 percent of the websites it surveys, a sample of more than 20 million sites.

[109] Niantic Spatial, "Thank You for Building with Us: The Next Chapter for 8th Wall," and 8th Wall, "Migration FAQ." From 28 February 2026, developers could no longer log in, edit their projects, or export them.

[110] "Awesome WebAssembly Languages," a community-maintained list that rates each language's WebAssembly support. Roughly sixty languages are rated as in production, stable, or usable.

[111] Dmitriy Semenkevich, "npm-registry-census," Zenodo, 27 August 2026, which counts 4,296,340 packages.

[112] Nabeel Al-Shamma and Thomas Nattestad, "Photoshop's journey to the web," web.dev, 26 October 2021; Autodesk, on the AutoCAD web app, which was built by compiling AutoCAD's existing C++ codebase, originally written in the 1980s, into WebAssembly.

[113] The Khronos Group, "Khronos Announces glTF Gaussian Splatting Extension," 3 February 2026. Contributors included Autodesk, Cesium and Bentley Systems, Esri, Huawei, Niantic Spatial, NVIDIA, and XGRIDS.

[114] Cesium, "Introducing 3D Gaussian Splats with Hierarchical Level of Detail Using 3D Tiles," 27 April 2026.

[115] W3Techs, "Content Management Systems Market Report," October 2026, which monitors 1,088 content management systems.

[116] GPS.gov, "GPS Accuracy," which puts GPS-enabled smartphones at about 4.9 meters under open sky.

[117] OpenStreetMap Wiki, "Stats."

[118] Overture Maps Foundation, "Overture Maps Foundation Reaches 50 Members as Industry Converges on Open Data to Ground AI," 13 July 2026.

[119] Niantic Spatial, "Locate," which reports more than one million VPS-activated locations, and Will Douglas Heaven, "How Pokémon Go Is Giving Delivery Robots an Inch-Perfect View of the World," *MIT Technology Review*, 10 March 2026, on the 30 billion images used to train its model.

[120] The Outernet Foundation, Placeframe repository.

[121] In the United Kingdom, the Online Safety Act requires "highly effective" age assurance from 25 July 2025 (Ofcom). In Australia, people under sixteen may not hold social media accounts from 10 December 2025

(eSafety Commissioner). In the United States, the Supreme Court upheld age verification for adult websites in *Free Speech Coalition v. Paxton* (June 2025).

[122] Regulation (EU) 2024/1183. Member states must offer a wallet by the end of 2026. It is free and voluntary. Recitals 22 and 60 cover pseudonyms.

[123] FIDO Alliance, "Five Billion Passkeys," 7 May 2026. The figure is the Alliance's own estimate.

[124] Shipped in Chrome 141 (30 September 2025) and Safari 26 (15 September 2025). Chrome for Developers, "Digital Credentials API: Secure and Private Identity on the Web."

[125] TechCrunch and *Variety*, 19 December 2025. Netflix acquired the company, and its services were wound down on 31 January 2026.

[126] Kim Cameron, "The Laws of Identity," Microsoft, May 2005.

[127] Raja Koduri, "Powering the Metaverse," Intel Newsroom, 14 December 2021.

[128] Max A. Cherney, "Mark Zuckerberg's Metaverse Will Require Computing Tech No One Knows How to Build," *Protocol*, 14 February 2022. Protocol has since closed, and the article is preserved by the Internet Archive.

[129] Apple Support, "Use Spatial Persona on Apple Vision Pro," which allows up to five participants.

[130] Ben Lang, "Facebook's Horizon Worlds Finally Hits (Kinda) Open Beta," *Road to VR*, 9 December 2021, on the 20-person limit at launch; *Road to VR*, 17 September 2025, on the rebuilt engine's "well over 100" concurrent users; and Meta Quest Help, "Set Up a Members-Only World in Meta Horizon Worlds," on the 25-visitor limit.

[131] VRChat, "May 22 2026 Developer Stream Q&A Summary."

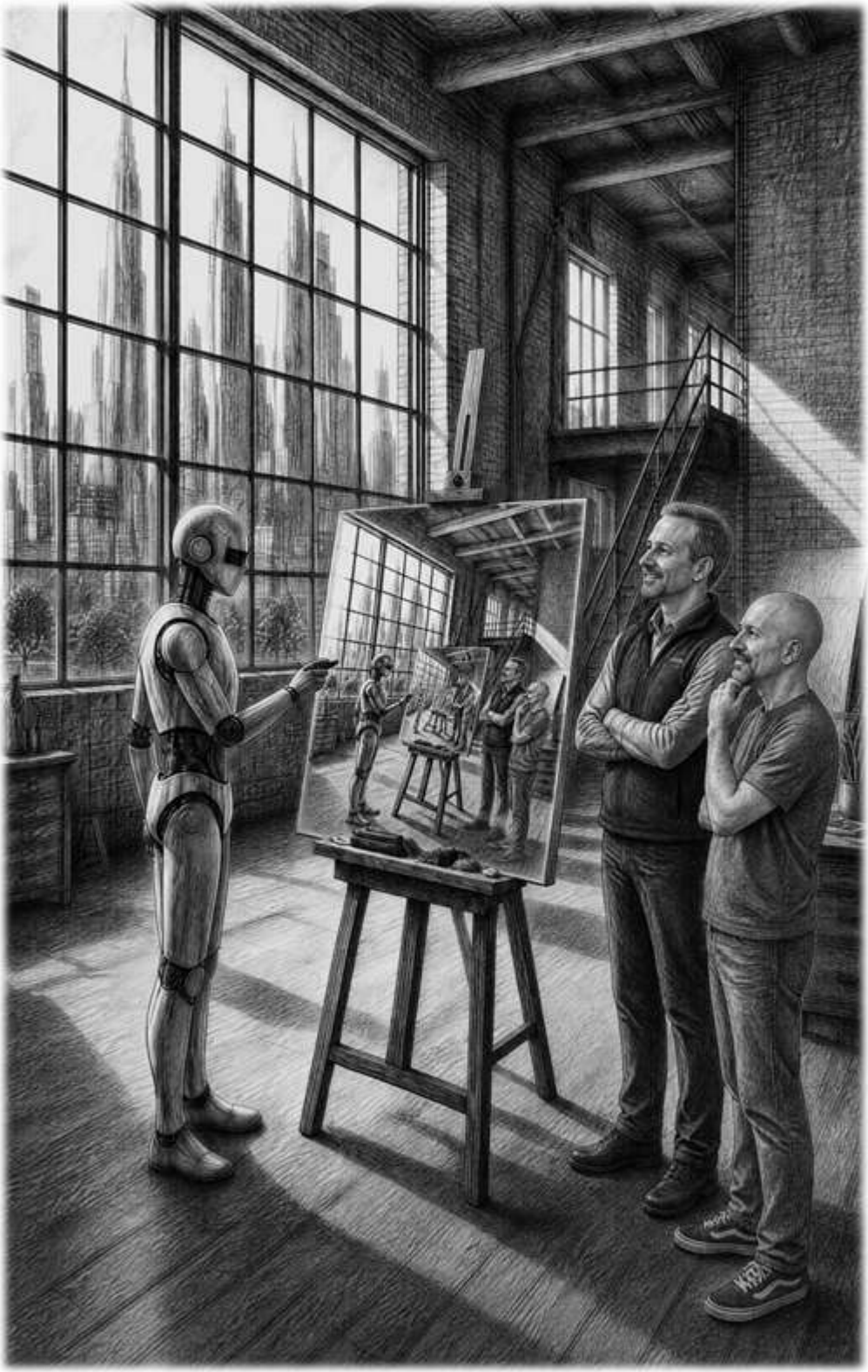
[132] Snap for Developers, "Connected Lenses" documentation for Spectacles. Sessions are colocated and limited to 64 participants.

[133] WHATWG, "Steering Group Agreement," and web.dev, "Interop 2026: Continuing to Improve the Web for Developers."

[134] *Los Angeles Times*, 14 September 1988; *TIME*, "Teaming Up Against Big Blue," September 1988; *The New York Times*, 20 September 1988; and *Chicago Tribune*, 20 November 1988. Later market shares are from Gartner and *PC World* annual vendor rankings.

[135] National Technology Transfer and Advancement Act of 1995, Pub. L. 104-113, §12(d), and Office of Management and Budget, Circular A-119, revised 22 January 2016.

[136] University of Rochester, "Open Metaverse Academic Alliance (OMAA)." Membership figure provided to the authors by the University of Rochester Center for Extended Reality (CXR), 2026.



About the Authors

Sean Mann

Sean Mann is CEO and co-founder of RP1, where he sets the company's vision and strategic direction, helping shape the next evolution of the open web. He asserts that the metaverse is fundamentally an infrastructure problem, not a product problem, and that enterprises will not adopt it unless the underlying foundation is open. That conviction drives his work with the Metaverse Standards Forum, where he serves on the Board of Directors. He helped launch the Forum's first initiative, the Open Metaverse Browser Initiative, and contributed RP1's open-source Sneeze browser engine and Rubidium browser as its foundation – designed to ensure that no single company, including RP1, controls the underlying technology.

Sean believes open standards are adopted only when industry, academia, and government align around the standards the next era of computing will require. He works with some of the world's largest technology companies, the W3C, and NIST to advance metaverse open standards across their platforms. With the University of Rochester, he helped establish the Open Metaverse Academic Alliance, now spanning 40+ universities to advance the research and develop the talent needed to build this open future.

Sean has spoken at the AR/VR Policy Conference and engages with members of Congress on the policies that will govern the web's next era of spatial computing and AI. His work is grounded in more than 25 years of taking new products from concept to market across medical devices, wellness, consumer electronics, fintech, and software. Named Innovator of the Year at the 2023 Polys WebXR Awards, Sean credits the recognition to the collective innovation and pioneering work of the RP1 team. He was also ranked among the Top 20 Metaverse Industry Founders by Metaverse Insider. A frequent speaker at AWE and SIGGRAPH, Sean has been interviewed on numerous podcasts and quoted by CNBC on spatial computing, with his work featured in *Forbes* and *VentureBeat*.

Dean Abramson

Dean Abramson is Chief Architect and co-founder of RP1, where he leads the innovation, research, design, and architecture behind the company's core technology. His central insight is that the metaverse isn't a destination or a product but the next evolution of the web – an open, connected space that anyone can build on and no one has to own.

Dean is obsessed with optimization. He has spent *0x23* years architecting and coding software systems that run at scale. As Chief Architect of Full Tilt Poker, one of the world's largest online poker platforms, he architected and built a system that supported more than 100,000 concurrent players. He used this experience to invent an entirely new architecture for networked, real-time, persistent, unsharded systems, in which every user shares one live world instead of being split across separate copies. That architecture became the basis for RP1's universal spatial fabric, which has demonstrated 100,000 concurrent high-fidelity avatars in a single shared space with full spatial audio and six degrees of freedom (6DoF) movement, an industry first.

His vision for an open metaverse led him to invent many of the concepts it depends on, including the multi-origin scene object model (SOM), which lets content from many independent sources coexist in one space, along with spatial fabrics, proximity-based browsing, and an open ecosystem of third-party services. He also invented the inverted scene graph technique that renders universe-sized worlds with millimeter precision using 32-bit graphics engines.

Dean invented, architected, and wrote code for: Sneeze, the world's first metaverse browser engine; Rubidium, the first open-source native metaverse browser; and Artemis, RP1's own native metaverse browser derived from Rubidium. Sneeze and Rubidium now form the open-source foundation of the Open Metaverse Browser Initiative at the Metaverse Standards Forum.

