Building Immersive Digital Twins: An Empirical Stakeholder Study of Barriers and Open Source Building Blocks in the European AR, VR, and IoT Ecosystem

Harsh Sudhir Shah[1], Özlem Mis[2], Bosen Zhou[3], Christian Zimmer[4], and Frank Feinbube[5]

Dietmar-Hopp-Allee 16 69190 Walldorf, Germany
harsh.shah01@sap.com, frank.feinbube@sap.com

Abstract. Immersive digital twins use virtual reality, augmented reality, and Internet of Things technologies to create interactive virtual replicas of physical assets and environments that users can enter and explore. Across Europe, experts report only limited progress toward immersive digital twins that are scalable, sovereign, and AI-augmented. The literature does not yet offer a stakeholder-grounded account of the barriers. This paper reports a reflexive thematic analysis of 11 semi-structured interviews with experts. The sample spans industrial manufacturing, automotive, telecommunications, public infrastructure, 2 platform or studio vendors, a research institute, a technology consultancy, 2 universities, and an industry association. The transcripts were coded collaboratively by 3 researchers who reached consensus through structured discussion. The qualitative analysis is corroborated by a Likert layer reported alongside the themes. The analysis identifies 8 master barrier themes. Interoperability across data formats and standards is the central constraint. Content creation is the dominant scaling bottleneck. The remaining themes cover hardware and platform volatility, the absence of a shared immersive user experience grammar, the role of artificial intelligence inside the runtime, trust and explainability limits, open source as a sovereignty question, and the demand for ready-to-run workflow scaffolding. A single-source observation concerns the erosion of open-source community review quality by AI-generated commits. The themes are translated into 7 candidate open-source building blocks that would relieve the common barriers. The contribution is empirical and forward looking. It is intended to guide open-source contribution priorities in the European immersive digital twin community.

Keywords: digital twins, immersive technologies, augmented and virtual reality, interoperability, open-source software, qualitative analysis

1 Introduction

Digital twins are virtual representations of physical objects, systems, or processes that are continuously updated with data from their physical counterparts

(Grieves, 2014; Grieves and Vickers, 2017). The concept has matured over the past decade across manufacturing, infrastructure, healthcare, and urban planning (Tao et al., 2019; Fuller et al., 2020). Kritzinger et al. (2018) observed that many systems described as digital twins are in fact digital models or digital shadows, where data flow is either absent or unidirectional, and that only a fraction realise the bidirectional integration that defines a true digital twin. Immersive digital twins extend this concept further. They are presented as navigable spatial environments that users enter through head-mounted displays or spatial computing devices (Slater and Sanchez-Vives, 2016). The combination of virtual reality (VR), augmented reality (AR), and Internet of Things (IoT) sensing forms a converging technology stack.

Despite this progress, scalable and sovereign immersive digital twins remain difficult to develop. Organisations rely on commercial engines such as Unity, and recent changes to their pricing and licensing have reduced confidence in this foundation. Platform deprecations by vendors have stranded earlier investments. Open-source alternatives exist, but they are often immature, poorly documented, or hard to adopt. Recent surveys of digital twin enabling technologies document technical gaps in standardisation, model heterogeneity, and data exchange across systems (Fuller et al., 2020; Lim et al., 2020). Less is known about which of these gaps block deployment in practice, seen from the perspective of the experts who develop and operate these systems.

This paper addresses 3 research questions:

RQ1. What are the dominant barriers to scalable immersive digital twin deployment across the European ecosystem?

RQ2. Which barriers are common across stakeholder roles, and which are role specific?

RQ3. Which open-source building blocks would materially relieve the common barriers?

We answer these questions through a reflexive thematic analysis of 11 semi-structured interviews with experts. The sample covers 11 organisations across the ecosystem and spans 4 stakeholder roles plus academic researchers. The analysis identifies 8 master barrier themes. We translate the themes into 7 candidate open-source building blocks. The contributions are twofold. First, an empirical barrier taxonomy grounded in stakeholder data. Second, a candidate building block agenda that maps the barriers to open-source targets. We also report a single-source observation on artificial intelligence (AI) generated commit erosion in open-source repositories, which we flag for future investigation.

This work takes as its premise that scalable immersive digital twins need an open foundation and cannot depend on proprietary platforms alone. To ground that premise in current practice, we examined barriers across the European immersive digital twin ecosystem rather than within a single application domain. The funding context and the resulting researcher position are disclosed in Section 3.6.

Section 2 reviews related work, Section 3 the research design, Section 4 the barrier themes with their Likert corroboration, Section 5 the building-block and research agenda, Section 6 the limitations, and Section 7 concludes.

2 Background and related work

2.1 Digital twins and their immersive extension

Grieves introduced the digital twin concept in 2003 as a physical product, a virtual product, and the connection between them (Grieves, 2014; Grieves and Vickers, 2017). Tao et al. (2019) extended this to a 5-dimension model that adds data and service dimensions and reported that no consensus on a unified modelling framework exists and that standardised connection protocols are still needed. Kritzinger et al. (2018) gave the most cited classification: a Digital Model has no automatic data exchange, a Digital Shadow has one-way flow from physical to digital, and a Digital Twin has bidirectional automatic flow. They found that most artefacts published as digital twins are in fact digital shadows. Trauer et al. (2020) add that a digital twin requires a representation of a physical system, bidirectional data exchange, and a connection across the product lifecycle.

Fuller et al. (2020) identified the absence of a unified model, the difficulty of operating heterogeneous systems, and the lack of standardisation as the most consequential challenges, alongside data, privacy, security, trust, and artificial intelligence integration. Lim et al. (2020) named virtual reality integration among 8 future perspectives, as a low-cost route to better visualisation and collaboration.

Immersive digital twins differ from conventional ones in how users engage with them. A conventional digital twin is monitored through a dashboard; an immersive digital twin is entered as a spatial environment. Slater and Sanchez-Vives (2016) explain that immersive virtual reality produces a place illusion and a plausibility illusion, which together make users behave as they would in the corresponding physical setting. Pirker et al. (2022), in a PRISMA review of 30 studies, found Industry 4.0 to be the dominant application domain and named interaction design, network optimisation, and hardware controls as open research areas. Doolani et al. (2020) reported that current head-mounted displays still face unsolved user experience problems, including haptics, latency, and bulk, so adoption depends on hardware improvement and lower-cost mobile alternatives.

2.2 Interoperability standards in the digital twin space

Open standards support immersive digital twins at several layers of the stack: scene description, asset delivery, and runtime access to devices. OpenUSD, originated by Pixar and now stewarded by the Alliance for OpenUSD, provides a hierarchical scene description and composition format that is widely used in industrial visualisation. The Alliance ratified the first OpenUSD Core Specification 1.0 in 2025 (Alliance for OpenUSD, 2025). The glTF 2.0 specification, maintained by the Khronos 3D Formats Working Group and adopted as ISO/IEC 12113:2022, provides a JSON-based transmission format for 3D assets that is optimised for real-time delivery and rendering (Khronos Group, 2021). The OpenXR specification is also maintained by Khronos. It provides a runtime application programming interface that abstracts across head-mounted displays and immersive devices, so that an application written once can target multiple vendor runtimes. Version 1.1, released in 2024, consolidated several interaction profiles and extensions into the core (Khronos Group, 2024).

Two further standards support the semantic and simulation layers of immersive digital twins. The Asset Administration Shell, defined in IEC 63278-1:2024, provides a structured metadata representation for industrial assets in the Industry 4.0 reference model. The shell exposes asset properties through modular submodels and supports interoperable identification across vendors (International Electrotechnical Commission, 2024). The Functional Mock-up Interface, originated by the MODELISAR consortium and now maintained by the Modelica Association, defines a tool-independent standard for the exchange of dynamic simulation models. Version 3.0, released in 2022 and updated in 2023, supports model exchange and co-simulation and names the next generation of digital twins among its target use cases (Modelica Association, 2022).

The presence of these standards does not by itself deliver interoperability. The Asset Administration Shell delivers syntactic interoperability, but semantic interoperability across heterogeneous models is still required, and recent work proposes ontology-based modelling of the shell to address this gap (Bousdekis, 2025). David et al. (2024) reported on a 2023 panel at the Annual Modeling and Simulation Conference that gathered strategic, technical, standardisation, and organisational perspectives on digital twin interoperability. The panel concluded that the state of practice limits interoperability to low levels and that higher levels are required for systems of digital twins to operate across vendor boundaries. The panel pointed to 3 emerging directions: co-simulation, standardisation, and infonomics. Trauer et al. (2020) reached a complementary conclusion through their case study. A consistent definitional foundation is a precondition for interoperability across the lifecycle stages they distinguish.

2.3 Artificial intelligence integration and the world model gap

Digital twins increasingly combine physics-based models with machine learning. Rasheed et al. (2020) framed this as hybrid analysis and modelling, and Kapteyn et al. (2021) gave a probabilistic foundation for predictive twins at scale. A known limit is that large language models can describe physical phenomena but cannot compute physical trajectories, the world model gap (Ha and Schmidhuber, 2018; LeCun, 2022). Recent world models begin to close it: V-JEPA 2 demonstrates prediction and planning from learned video (Assran et al., 2025), and Cosmos generates physics-based simulation of industrial environments (NVIDIA, 2025). This gap underpins Theme 5, where artificial intelligence moves from helper to runtime infrastructure.

2.4 Trust and explainability

Holzinger et al. (2019) distinguished explainability, a property of a system, from causability, the measurable quality of an explanation in supporting human understanding, and argued that trust in safety-critical artificial intelligence requires both. Liao and Varshney (2022) framed explainable artificial intelligence as a function of the user's task and mental model. These concerns matter for digital twins because agentic systems inside twins are beginning to act, where earlier systems only reported observations.

2.5 Open source as ecosystem strategy

Lerner and Tirole (2002) offered the foundational economic analysis of open-source software. They argued that the contributions of individual developers can be explained by labour economics and the literature on career concerns, where signalling and peer recognition motivate participation. The European Union has framed data sovereignty and open infrastructure as strategic concerns, notably through initiatives such as Gaia-X, which proposes a federated data spaces architecture as a sovereign alternative to non-European hyperscale platforms (Gaia-X European Association for Data and Cloud, 2025). The findings reported in this paper revisit each of these strands and raise a novel sub-finding on AI-generated commit quality that is not present in the existing open-source economics literature.

3 Methodology

3.1 Research design

We adopted reflexive thematic analysis as the primary method (Braun and Clarke, 2006, 2019), which suits the study of meaning making across a diverse sample and treats the researchers' interpretive engagement as a resource. We complemented it with a small layer of 5 Likert rating questions, used to surface how experts position themselves on key issues and to prompt their reasoning, rather than as a quantitative measure. The Likert layer is used for qualitative corroboration only, not for inferential claims, and is reported in Section 4 alongside the themes it supports.

3.2 Sample and recruitment

We interviewed 11 external experts through purposive sampling across the European AR, VR, and digital twin ecosystem. By experts we mean professionals and academic researchers with direct experience of building, deploying, or governing immersive digital twins. The sample was designed to span 4 stakeholder roles, together with academic researchers. Producers are vendors and research institutes that develop digital twin solutions. Consumers are organisations that adopt them. An intermediary connects builders and buyers. A governance role represents the industry view. Several experts also hold academic or research positions, marked as researcher in Table 1, a distinction that is relevant to the open-source findings reported in Section 4.7.

Sample size is consistent with saturation guidance. Guest et al. (2006) found basic themes emerged within 6 interviews and saturation by 12 in a homogeneous sample; Hennink et al. (2017) reported code saturation at 9 and meaning saturation at 16 to 24. The present sample of 11 is deliberately heterogeneous, favouring breadth over depth on any single role, so the analysis claims code-level rather than meaning saturation, noted as a limitation in Section 6.

Table 1 summarises the sample. All identities are anonymised in accordance with the project's data protection plan. Role descriptors and sector tags support analytical generalisation (Yin, 2018).

Table 1: Expert sample by role, sector, and stakeholder role (n = 11).

ID	Role	Sector	Stakeholder role
I01	Head of Foundational Technologies	Industrial Manufacturing	Consumer and Producer
I02	Immersive Technology Lead	Industry Association	Governance
I03	Chief Executive Officer	Virtual Reality Collaboration Platform	Producer
I04	Senior Scientist	Research Institute	Producer and Researcher
I05	Professor of Computational Design	Architecture University	Consumer and Researcher
I06	Chief Executive Officer	Extended Reality (XR) Studio	Producer
I07	Reality Capture Lead	Public Infrastructure (Rail and Geospatial)	Consumer
I08	Portfolio Lead, Web, Mobile, and Extended Reality	Telecommunications	Intermediary
I09	Professor of Immersive Systems	Technical University	Researcher
I10	Technical Architect	Automotive Original Equipment Manufacturer (OEM)	Consumer
I11	Artificial Intelligence Laboratory Leader	Technology Consulting	Producer and Researcher

3.3 Interview instrument

The instrument was a semi-structured guide of approximately 60 minutes, organised in 5 sections. Section 1 covered role context and technology stack. Section 2 probed development challenges, capability gaps, and interoperability. Section 3 elicited an ideal open-source ecosystem vision. Section 4 covered specific technology areas, including immersive devices, the Asset Administration Shell, and artificial intelligence. Section 5 included 5 Likert rating questions on a 1 to 5 scale, a long-horizon vision question, and a magic wand question that elicited the single most important fix the expert would make. All interviews were conducted in English and recorded with informed consent. Nine interviews were transcribed verbatim, and two were captured as detailed summary notes rather than full transcripts. Two note takers were present at all 11 interviews, in addition to the audio recordings.

3.4 Coding procedure

The 11 transcripts were coded collaboratively by 3 researchers. The procedure drew on the reflexive thematic analysis approach of Braun and Clarke (2006, 2019). The team familiarised itself with the data through full readings of each transcript. Each researcher generated an initial set of codes independently. The team then met to consolidate the code inventory, discuss disagreements, and reach consensus through structured discussion. Codes were clustered into sub-themes and then into master themes. The theme structure was added to, modified, and reorganised through multiple rounds of refinement, and the final structure was reached by consensus across the 3 researchers.

This procedure follows Braun and Clarke's (2019) position that reflexive thematic analysis does not require inter-coder reliability metrics; reliability follows from analytical depth and reflexive engagement. The 3-researcher consensus goes beyond a single-coder design and provides an audit trail.

3.5 Trustworthiness

Trustworthiness was addressed using the 4 criteria of Lincoln and Guba (1985). Credibility rests on the consensus coding procedure across 3 researchers. Transferability rests on the role-stratified sample in Table 1. Dependability rests on the audit trail of the consensus coding discussions. Confirmability rests on full anonymisation before coding and on retention of the original transcripts under the project's data protection plan. Experts provided informed consent for the recording, transcription, and publication of anonymised excerpts.

3.6 Reflexivity statement

This study was conducted within Immersive Data Center Management (IDCM), a project of the European Union funded ApeiroRA initiative (Apeiro Reference Architecture), which develops open-source reusable components for a sovereign cloud-edge infrastructure. The intent of IDCM is to make scalable immersive digital twins easier to develop through open-source contributions. The interviewers and coders are affiliated with project teams that have a stake in the success of an open ecosystem. We mitigated this position through 3 measures. We used open prompts in the interview guide rather than leading questions where possible. We used 3-researcher consensus coding rather than single-coder analysis. We disclose the sponsor and the project framing here. The identities of the experts are anonymised and will remain so.

4 Findings: Eight master barrier themes

The analysis identified 8 master themes. Table 2 reports how many of the 11 experts raised each theme. Five themes appear in all 11 interviews, and every theme was raised by at least 6 of the 11 experts.

Table 2: Master barrier themes and the number of experts (of 11) who raised each.

ID	Theme	Coverage
T1	Interoperability as the central binding constraint	11 of 11
T2	Content creation as the dominant scaling bottleneck	10 of 11
T3	Hardware and platform volatility	11 of 11
T4	The missing shared grammar in immersive user experience	11 of 11
T5	Artificial intelligence from helper to runtime infrastructure	11 of 11
T6	Trust and the limits of autonomous action	9 of 11
T7	Open source as a sovereignty and resilience question	11 of 11
T8	Workflow scaffolding and developer experience	6 of 11 strong, 9 of 11 with partial mention

Each theme is summarised in prose, with anonymised expert views attributed by identifier (I01 through I11) and role. Representative verbatim quotations are provided as supplementary material.

The interview instrument also included 5 Likert rating questions on a 1 to 5 scale. Ten of the 11 experts gave at least 1 numerical rating, and 1 expert (I11) answered the items qualitatively. A small number of items also drew a qualitative rather than numerical answer from individual experts, so the item-level counts differ. Table 3 reports the means. The Likert layer is a corroborating signal only. The per-interview ratings are provided as supplementary material. Each item is discussed below with the theme it supports.

Table 3: Means of the 5 Likert items on a 1 to 5 scale, with item-level response counts.

Item	Description	Mean (n)
L1	Importance of open-source availability	3.1 (n = 9)
L2	Current availability of open-source digital twin components meets needs	2.4 (n = 9)
L3	Value of pre-built customisable blueprints	4.3 (n = 10)
L4	Importance of AI and automation features	4.3 (n = 10)
L5	Criticality of community support for open-source adoption	5.0 (n = 9)

4.1 Theme 1: Interoperability as the central binding constraint

All 11 experts identified interoperability as the dominant barrier. The barrier appears in multiple forms. Engineering computer-aided design files do not convert cleanly to real-time rendering meshes. Time series data from Internet of Things sensors, geometric models, and structured metadata are held in separate systems that do not share a common semantic vocabulary. Brownfield retrofit onto existing information technology estates is recurrent in the manufacturing, automotive, and public infrastructure transcripts. The growing number of formats across OpenUSD, glTF, the Asset Administration Shell, and the Functional Mock-up Interface multiplies the integration cost. The industrial manufacturing expert noted that export control rules can also affect interoperability, because 3D geometry can fall under technology transfer regulations when it is shared across borders within a multinational.

The artificial intelligence laboratory leader proposed an ontology layer as the central missing building block. The field has good standards for identification and authorisation but lacks a shared semantic foundation. This proposal informs the open-source agenda in Section 5.

The dominance of this theme is corroborated by an independent prompt in the instrument. 8 of 11 experts named interoperability or data integration as the magic wand fix they would make today. The industrial manufacturing expert made the wish explicit: "Interoperability. If we can easily plug things together with a magic way, just throw a line and information flows from one position to the next one." The immersive technology lead from the industry association articulated the same desire as removing the borders between data, data formats, and the walls between different sectors and companies. The convergence between thematic coding and the magic wand prompt is the strongest single signal in the dataset.

4.2 Theme 2: Content creation as the dominant scaling bottleneck

Ten of 11 experts framed content creation as a scaling bottleneck. Bespoke geometry, scene composition, lighting, and interaction setup are required for each new use case. Three reinforcing factors appear. First, accuracy mismatch between engineering computer-aided design tools and visualisation tools requires manual conversion. Second, some assets cannot be acquired at any reasonable cost. Third, the surrounding environment around the geometric core is itself a content creation problem.

AI-assisted generation, including scene generation tools and Gaussian splatting-based scanning, appears across several interviews as a partial mitigation but not as a standardised solution.

4.3 Theme 3: Hardware and platform volatility

All 11 experts identified hardware diversity and vendor volatility as a sustained challenge. Devices in use include the Meta Quest family, Pico, the Apple Vision Pro, the Microsoft HoloLens 2, the Varjo headsets, and emerging smart glasses. Battery life, weight, and ergonomic limits restrict deployment in industrial settings, especially where workers move continuously through hazardous spaces. The discontinuation of the Microsoft HoloLens line and the deprecation of certain platform software development kits were cited as concrete events that stranded prior investments.

Several experts described the lack of mainstream lightweight see-through augmented reality glasses as the next inflection point that the field is waiting for. The automotive expert noted the graphics processing unit and ray tracing constraints of standalone head-mounted displays.

4.4 Theme 4: The missing shared grammar in immersive user experience

All 11 experts referenced user experience inconsistency. The same controller has different button mappings across different applications on the same headset. There are at least 2 design philosophies in active circulation. One school treats the immersive environment as a spatial extension of a personal computer interface, with menus and pointers. The other school treats the immersive environment as a mirror of the physical world, where a user walks to a switch and presses it. There is no consensus on which approach is preferable for which task. Onboarding friction discourages occasional users. A 2025 systematic review of augmented reality interaction design reached the same conclusion, finding that user experience evaluation methods and interaction conventions are not yet established across devices (Hughes and Karwowski, 2025).

The automotive expert offered a counter view. Integration must be solved first because user experience is downstream of the integration foundation. This view is reflected in the engineering implication discussion in Section 5.

4.5 Theme 5: Artificial intelligence from helper to runtime infrastructure

All 11 experts discussed artificial intelligence integration. The discussion clusters into 3 distinct sub-patterns at different levels of maturity.

In the first pattern, artificial intelligence is a helper. It assists with scene generation, format translation, or content authoring. This use is established and operational across the sample. The industrial manufacturing expert described using generative artificial intelligence to design the visual scene around a computer-aided design file. The architecture university expert described pattern recognition and digital model recovery as a strong domain for current artificial intelligence systems.

In the second pattern, artificial intelligence is a copilot across heterogeneous data. It retrieves relevant information when an operator asks a question. It performs root cause analysis across systems that would otherwise require manual aggregation. The industrial manufacturing expert described an artificial intelligence assistant that brings answers from the digital twin's single source of truth. The automotive expert described a factory copilot with a large language model chat interface and agentic workflows at proof-of-concept stage.

In the third pattern, artificial intelligence is part of the runtime infrastructure. Multi-agent systems coordinate remote factory management. Agents communicate with each other and feed back to a human operator. The technology consulting expert described one such system in their research laboratory. The industry association expert noted that the same pattern is being explored by their member organisations, often combined with augmented reality avatars.

Experts also pointed to the world model gap introduced in Section 2.3. Language models can describe physical phenomena but cannot compute trajectories, so they cannot simulate.

Two practical learnings follow from this pattern of evidence. First, hybrid digital twins that combine physics-based models with machine learning are the engineering answer that experts and the literature both converge on (Rasheed et al., 2020; Kapteyn et al., 2021). The scientist at the research institute described this combination as already in use because physics-based models alone are too slow for some decisions. Second, trust constraints become more demanding as artificial intelligence moves from helper to infrastructure. This pattern points ahead to Theme 6 and is revisited in the discussion section. The Likert layer corroborates the importance of this theme. The importance of AI and automation features (L4) returned a mean of 4.3.

4.6 Theme 6: Trust and the limits of autonomous action

Nine of 11 experts referenced trust as a limit on adoption. Trust grows through repeated verification. Model architecture matters less than a track record of accurate results.

The automotive expert ranked accuracy first and the ability to follow how the result was derived second among trust factors. Several experts cited hallucination risk as a concern, especially for safety-adjacent uses.

A deeper concern was raised by the telecommunications expert and echoed by the technology consulting expert. The hardest digital twin problems may be socio-organisational rather than technical because computers cannot represent social structures.

This sub-finding does not appear strongly in the existing trust and explainability literature on digital twins.

4.7 Theme 7: Open source as a sovereignty and resilience question

All 11 experts discussed open-source software, though their positions differ. Open source is treated as a hedge against vendor lock-in by some experts and as a sovereignty question by others. The Unity engine pricing change of 2023, which the vendor later rescinded in 2024 after community objection, and the deprecation of the Meta avatar software development kit were cited as concrete events that motivate caution about proprietary platforms. A vendor-funded community-maintained governance pattern was discussed by the architecture university expert, who cited the Ultimaker and Cura ecosystem as a working precedent.

A single-source sub-finding was raised by the virtual reality platform expert. AI-generated commits in open-source repositories have begun to flood community projects with low quality contributions that human reviewers cannot keep up with.

This observation comes from a single expert, and we do not present it as a saturated theme finding. We note that it matches a pattern now documented across open-source projects, including the closure of the curl bug bounty programme in early 2026 after a surge of AI-generated submissions (Stenberg, 2026). It does not appear in the existing open-source economics literature (Lerner and Tirole, 2002). It is discussed as a candidate future research direction in Section 5.

The Likert layer corroborates the divided stance. Importance of open-source availability (L1, mean 3.1) was polarised: the 2 commercial vendors rated it 1 and the 3 research or academic experts rated it 5. Current availability (L2) returned the lowest mean, 2.4, the strongest signal that supply lags need, while community support (L5) returned the highest, 5.0. One dissenting view, from the technical university expert, was that AI-assisted documentation may eventually reduce dependence on community size.

4.8 Theme 8: Workflow scaffolding and developer experience

Six experts raised workflow scaffolding as a primary need. Three more raised it partially. The pattern is consistent. The need goes past assets and libraries. Experts want pre-configured workflows, drag-and-drop primitives, and ready-to-run sample setups.

The automotive expert prioritised standardisation in the integration foundation over user experience polish. The Likert layer supports this theme. The value of pre-built customisable blueprints (L3) returned a mean of 4.3, among the highest of the 5 items.

4.9 Common and role-specific barriers

The barriers separate into common and role-specific, which addresses RQ2. Interoperability (T1), hardware and platform volatility (T3), and the missing interaction grammar (T4) were raised by all 11 experts. Role-specific emphases also appeared. The value placed on open source (L1) split by role: the 2 commercial vendors rated it lowest and the 3 academic or research experts highest. Content creation cost (T2) was most acute for the studio and producer roles. Brownfield integration weighed most on the consumer roles in manufacturing, automotive, and rail. The intermediary and governance roles framed the hardest barriers as organisational rather than technical.

5 Discussion

Existing surveys derive digital twin challenges from the literature. Fuller et al. (2020) name the unified model, heterogeneity, and standardisation, and David et al. (2024) rank interoperability maturity. The present study is stakeholder-derived rather than literature-derived. It surfaces 3 barriers that the technical surveys cover only lightly: the content creation bottleneck, the missing interaction grammar, and the demand for ready-to-run workflow scaffolding.

5.1 An open-source building block agenda

We use the term *building block* to refer to a small, well-defined, reusable, open-source component that addresses a specific barrier. A building block composes with others into a working system. It is not itself a platform or a finished product. Small, composable components are easier to maintain and combine than monolithic platforms. We propose that the immersive digital twin community adopt this principle.

Table 4 reframes the 8 barrier themes as a candidate agenda of 7 open-source building blocks. The mapping is not one to one: trust (Theme 6) is treated as a cross-cutting concern addressed through provenance metadata in the artificial intelligence building block, and sovereignty as a property of ecosystem governance rather than a block in its own right. Each row names the closest existing precedent, to show that each block extends existing open work rather than starting from nothing.

Table 4: Mapping of the 8 barrier themes to 7 candidate open-source building blocks, with closest precedent and contribution dimension. Trust (T6) is a cross-cutting concern addressed within the artificial intelligence block rather than a separate row.

Barrier theme	Proposed open-source building block	Closest precedent	Contribution
T1 Interoperability	An open ontology and connector layer that bridges OpenUSD, glTF, the Asset Administration Shell, and the Functional Mock-up Interface through a shared semantic foundation	Eclipse BaSyx, the open Asset Administration Shell middleware whose Resource Description Framework (RDF) support is still experimental; OpenUSD with Asset Administration Shell submodels; ontology modelling of the shell (Bousdekis, 2025); the Gaia-X data spaces architecture	A reference implementation that supplies the semantic connector and provenance layer left experimental in current tooling, addressing the low level of interoperability identified by David et al. (2024)
T2 Content creation	An open computer-aided design to extended reality conversion toolchain with provenance preservation	Mature computer-aided design to OpenUSD to glTF converters; Khronos glTF sample models; Gaussian splatting reality capture	A standardised, openly licensed pipeline that preserves the link from source to mesh, which current converters drop, and reduces per-project manual effort
T3 Hardware volatility	A packaging and capability-tier layer above OpenXR and WebXR for portable immersive application delivery	OpenXR and the Monado open runtime already provide runtime portability; WebXR covers the browser	A build and capability-abstraction layer above the runtime portability that OpenXR and Monado already deliver, so that one application package deploys across device tiers and the browser

continued on next page

Table 4 (continued)

Barrier theme	Proposed open-source building block	Closest precedent	Contribution
T4 User experience	A reusable interaction primitive library with a public, engine-neutral conformance suite	Open interaction toolkits exist but are engine-bound, such as MRTK3 and Godot XR Tools; OpenXR 1.1 interaction profiles	The first engine-neutral and conformance-tested set of interaction primitives, since existing toolkits are engine-specific and OpenXR conformance covers the runtime rather than interaction grammar
T5 Artificial intelligence integration	An open agent runtime that brokers between language models, physics engines, and ontologies, with embedded provenance and derivability metadata	LangGraph, which is open source, for coordination; open physics through MuJoCo and Isaac; world models such as V-JEPA 2 (Assran et al., 2025) and Cosmos (NVIDIA, 2025); NVIDIA Omniverse libraries are partly open but remain platform dependent	A neutral runtime that composes language and physics models over the T1 ontology and addresses the world model gap, with built-in trust artefacts
T7 Open source ecosystem	A vendor-funded community-maintained governance pattern with tooling to triage AI-generated commits	Cura and Ultimaker, the Blender Foundation, and the Eclipse Foundation for governance; an open triage tool for AI-generated commits is still nascent	An established governance template plus a novel open tool for the single-source sub-finding in Theme 7, now corroborated by 2025 and 2026 reports (Stenberg, 2026)
T8 Workflow scaffolding	A platform-neutral reference architecture with drag-and-drop primitives and ready-to-run sample workflows	NVIDIA Omniverse Blueprints are open reference workflows but tied to a proprietary platform; Hugging Face Spaces builds on the open Gradio library as a hosted service	An open and platform-neutral scaffold that matches the empirical demand identified in Theme 8

The strongest novelty sits in 3 rows: the interoperability block names the ontology layer as an open target, building on experimental RDF work in Asset

Administration Shell tooling; the open-source ecosystem block calls for an open tool to triage AI-generated commits; and the workflow scaffolding block captures a demand the literature under-weights. The cross-runtime block overlaps with what OpenXR and Monado already provide, so it is framed narrowly as a packaging and capability layer above them.

Several experts framed open source as a question of European digital sovereignty, citing export controls on cross-border 3D geometry and dependence on hardware vendors outside Europe. Gaia-X is the closest European policy precedent (Gaia-X European Association for Data and Cloud, 2025). The findings suggest the immersive digital twin layer needs an equivalent set of open building blocks if European actors are to stay resilient against vendor exits and pricing changes.

5.2 Latent themes and a research agenda

Two latent themes cut across the master themes. The brownfield problem is that industrial immersive digital twins are retrofitted onto existing IT estates with uneven data quality and weak interfaces, which shapes the interoperability and content-creation blocks. The socio-organisational ceiling, raised by 2 experts, is that the hardest barriers are misaligned incentives and norms that cannot be represented as data, which software blocks alone cannot resolve. Both warrant dedicated study.

The analysis points to 5 research directions. First, an empirical study of AI-generated commit erosion, combining commit-log analysis with maintainer interviews, since no peer-reviewed study yet documents it (Stenberg, 2026). Second, a prototype of the ontology connector layer bridging OpenUSD, glTF, and the Asset Administration Shell with provenance, measured against the manual baseline experts described. Third, a controlled comparison of hybrid physics plus machine learning twins against language-model-only coordination. Fourth, a multi-site evaluation of the workflow scaffold, measuring time to a first working twin against building from scratch. Fifth, a cross-regional replication beyond Europe, treating the 8-theme taxonomy and 7-block agenda as hypotheses rather than conclusions. The first asks whether the ecosystem can sustain the blocks, the next three whether each block delivers, and the fifth whether the taxonomy generalises.

6 Limitations

Five limitations to validity are noted. First, the sample is European and reflects a regional perspective. Replication in North American, East Asian, or Indian ecosystems would strengthen the external validity of the barrier taxonomy. Second, the interviewers and the coders are affiliated with the funding initiative, which has an open-source mandate. This creates a research interest position, disclosed in Section 3.6. The open prompt structure of the instrument and the 3-researcher consensus coding procedure mitigate it. Third, the interview instrument explicitly framed Section 3 around an ideal open-source ecosystem. This is a leading framing. It may have biased experts toward open-source language even where their substantive views were more agnostic. A subsequent instrument design should test substantive views with neutral phrasing. Fourth, the sample is purposive rather than random. The findings support analytical generalisation (Yin, 2018) rather than statistical generalisation. The sample reaches code saturation but not meaning saturation, as noted in Section 3.2. Fifth, the Likert layer is descriptive and not powered for inferential statistics. Its function is corroborative, and a small number of items drew qualitative rather than numerical answers, so item-level counts differ.

Two of the 11 interviews were recorded as detailed summary notes rather than verbatim transcripts, which limits the depth of quotation drawn from them. A further limitation is that the digital twin and immersive technology field is moving quickly. The findings reflect the state of practice at the time of data collection. We recommend that the barrier taxonomy be revisited at intervals of approximately 18 months.

7 Conclusion and future work

We reported a reflexive thematic analysis of 11 semi-structured interviews with experts in the European AR, VR, and Internet of Things digital twin ecosystem. The transcripts were coded collaboratively by 3 researchers who reached consensus on the theme structure. The analysis identifies 8 master barrier themes. Five appear across all 11 interviews. We corroborate the findings with a Likert layer reported alongside the themes.

The contributions are twofold. First, an empirical barrier taxonomy grounded in stakeholder data. Second, a candidate agenda of 7 open-source building blocks that map the barriers to specific contribution targets. We additionally flag a single-source observation on AI-generated commit erosion in open-source repositories as a direction for future work.

Future work spans 5 lines set out in Section 5.2: AI-commit erosion, a prototype ontology connector layer, a comparison of hybrid and language-model-only twins, a multi-site evaluation of the workflow scaffold, and a cross-regional replication. Together they convert this empirical contribution into a coherent research programme.

Acknowledgements

This research work was carried out as part of the project “Apeiro Reference Architecture”, in short ApeiroRA, under the funding ID 13IPC007, financed by the European Union – NextGeneartionEU and funded by the German Federal Ministry for Economic Affairs and Energy. We thank the 11 anonymous experts who contributed their time and expertise.

References

References

Alliance for OpenUSD (2025). OpenUSD core specification, version 1.0. Joint Development Foundation Projects, LLC.

Assran, M., Bardes, A., Fan, D., Garrido, Q., Howes, R., Muckley, M., . . . Ballas, N. (2025). V-JEPA 2: self-supervised video models enable understanding, prediction and planning.

Bousdekis, A. (2025). The semantic asset administration shell: ontology modelling of digital twins for sustainable interoperability in Industry 4.0. Intelligent Decision Technologies, 19(3), 1783–1806.

Braun, V., and Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3(2), 77–101.

Braun, V., and Clarke, V. (2019). Reflecting on reflexive thematic analysis. Qualitative Research in Sport, Exercise and Health, 11(4), 589–597.

David, I., Shao, G., Gomes, C., Tilbury, D., and Zarkout, B. (2024). Interoperability of digital twins: challenges, success factors, and future research directions. In: Margaria, T., and Steffen, B. (eds.), Leveraging Applications of Formal Methods, Verification and Validation. Lecture Notes in Computer Science, vol 15223, Springer, Cham.

Doolani, S., Wessels, C., Kanal, V., Sevastopoulos, C., Jaiswal, A., Nambiappan, H., and Makedon, F. (2020). A review of extended reality (XR) technologies for manufacturing training. Technologies, 8(4), 77.

Fuller, A., Fan, Z., Day, C., and Barlow, C. (2020). Digital twin: enabling technologies, challenges and open research. IEEE Access, 8, 108952–108971.

Gaia-X European Association for Data and Cloud (2025). Gaia-X architecture document, release 25.11.

Grieves, M. (2014). Digital twin: manufacturing excellence through virtual factory replication. White Paper, Florida Institute of Technology.

Grieves, M., and Vickers, J. (2017). Digital twin: mitigating unpredictable, undesirable emergent behavior in complex systems. In: Kahlen, F. J., Flumerfelt, S., and Alves, A. (eds.), Transdisciplinary Perspectives on Complex Systems, Springer, pp. 85–113.

Guest, G., Bunce, A., and Johnson, L. (2006). How many interviews are enough? An experiment with data saturation and variability. Field Methods, 18(1), 59–82.

Ha, D., and Schmidhuber, J. (2018). World models.

Hennink, M. M., Kaiser, B. N., and Marconi, V. C. (2017). Code saturation versus meaning saturation: how many interviews are enough? Qualitative Health Research, 27(4), 591–608.

Holzinger, A., Langs, G., Denk, H., Zatloukal, K., and Mueller, H. (2019). Causability and explainability of artificial intelligence in medicine. WIREs Data Mining and Knowledge Discovery, 9(4), e1312.

Hughes, C. L., and Karwowski, W. (2025). Evaluating interaction design and user experience in augmented reality: a systematic review. Frontiers in Virtual Reality, 6, 1710161.

International Electrotechnical Commission (2024). IEC 63278-1:2024 Asset Administration Shell for industrial applications, Part 1: Asset Administration Shell structure.

Kapteyn, M. G., Pretorius, J. V. R., and Willcox, K. E. (2021). A probabilistic graphical model foundation for enabling predictive digital twins at scale. Nature Computational Science, 1, 337–347.

Khronos Group (2021). glTF 2.0 specification, version 2.0.1 (ISO/IEC 12113:2022).

Khronos Group (2024). OpenXR specification, version 1.1. Khronos OpenXR Working Group.

Kritzinger, W., Karner, M., Traar, G., Henjes, J., and Sihn, W. (2018). Digital twin in manufacturing: a categorical literature review and classification. IFAC PapersOnLine, 51(11), 1016–1022.

LeCun, Y. (2022). A path towards autonomous machine intelligence, version 0.9.2. OpenReview.

Lerner, J., and Tirole, J. (2002). Some simple economics of open source. Journal of Industrial Economics, 50(2), 197–234.

Liao, Q. V., and Varshney, K. R. (2022). Human centered explainable AI: from algorithms to user experiences.

Lim, K. Y. H., Zheng, P., and Chen, C. H. (2020). A state of the art survey of digital twin: techniques, engineering product lifecycle management and business innovation perspectives. Journal of Intelligent Manufacturing, 31, 1313–1337.

Lincoln, Y. S., and Guba, E. G. (1985). Naturalistic Inquiry. Sage Publications.

Modelica Association (2022). Functional Mock-up Interface specification, version 3.0.

NVIDIA (2025). Cosmos world foundation model platform for physical AI.

Pirker, J., Loria, E., Safikhani, S., Künz, A., and Rosmann, S. (2022). Immersive virtual reality for virtual and digital twins: a literature review to identify state of the art and perspectives. Proceedings of the 2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW), 114–120.

Rasheed, A., San, O., and Kvamsdal, T. (2020). Digital twin: values, challenges and enablers from a modeling perspective. IEEE Access, 8, 21980–22012.

Slater, M., and Sanchez-Vives, M. V. (2016). Enhancing our lives with immersive virtual reality. Frontiers in Robotics and AI, 3, 74.

Stenberg, D. (2026). The end of the curl bug-bounty. daniel.haxx.se, 26 January 2026. https://daniel.haxx.se/blog/2026/01/26/the-end-of-the-curl-bug-bounty/

Tao, F., Zhang, H., Liu, A., and Nee, A. Y. C. (2019). Digital twin in industry: state of the art. IEEE Transactions on Industrial Informatics, 15(4), 2405–2415.

Trauer, J., Schweigert-Recksiek, S., Engel, C., Spreitzer, K., and Zimmermann, M. (2020). What is a digital twin? Definitions and insights from an industrial case study in technical product development. Proceedings of the Design Society: DESIGN Conference, 1, 757–766.

Yin, R. K. (2018). Case study research and applications: design and methods, sixth edition. Sage Publications.