PHANTOMATRIX: A Transfer Model for Extended Reality Systems

Armin Grasnick[*]

IU International University of Applied Sciences Germany

Abstract: Extended Reality (XR) systems mediate event fields through technical transformations that intervene between physical events and their rendered representation. This article formalizes these transformations within the PHANTOMATRIX framework, which models XR systems as composed mappings between primary event fields, phantom representations, and secondary event fields. Capture, transformation, and display are defined as sequential stages governed by mode-specific and system-imposed constraints. By expressing XR configurations as transformation chains, PHANTOMATRIX provides a formal basis for analyzing how technical constraints shape mediated environments and how different system architectures can be represented within a unified compositional structure. This perspective enables XR systems to be examined in terms of stage-specific constraints that influence mediated experience. Such constraints provide a basis for analyzing user experience effects in XR systems.

1. INTRODUCTION

Formal Context

Extended Reality (XR) technologies are studied and applied across domains, ranging from entertainment and gaming [1] to healthcare [2, 3], workplace contexts [4], and education [5, 6]. These technologies, sometimes referred to as digital reality [7], rely on technical systems that capture, transform, and present multimodal sensory data. XR systems have been investigated with respect to their influence on perceptual and affective processes [8-10]. Various physical and perceptual factors influence the performance of VR systems [11].

Prior research approaches XR systems from different perspectives. Conceptual models classify configurations along the reality-virtuality spectrum [12], while other studies focus on specific technical processes such as latency and motion-to-photon dynamics [13, 14]. These contributions provide insights into individual layers of XR systems, yet they do not articulate an integrated formal representation that systematically relates capture, transformation, and presentation processes within a single framework.

Empirical research indicates that perceptual and neural processing can differ between direct and mediated interaction contexts [15]. Such findings suggest that technical mediation systematically influences perceptual processing. XR systems interpose technical operations between physical events and their mediated presentation. These operations form a transformation sequence that is typically addressed through domain-specific models rather than within a unified formalism.

While prior work addresses technical aspects of XR systems (*e.g.*, latency, sensing, display fidelity) and experiential outcomes (*e.g.*, perception, affect, engagement) in parallel, the relation between system architecture and experiential design decisions often remains implicit. Technical modeling typically focuses on isolated components or performance metrics, whereas user experience research evaluates perceptual and affective responses at the system level. A formal framework that localizes constraints within the mediation chain can help bridge this gap by mapping experiential effects to specific architectural stages.

Framework Definition

We introduce PHANTOMATRIX[1], a formal structure based on transfer functions, multimodal event fields, and phantom representations in XR systems. An earlier version of the PHANTOMATRIX framework was published as a preprint [16]. The present article revises the conceptual orientation of the model and provides a more rigorous formalization of its transformation behavior. The previous emphasis on perceptual convergence has been removed in favor of a strictly formal representation of the transformation chain relating to physical events, technical mediation, and mediated presentation.

The PHANTOMATRIX framework formalizes XR systems as a transformation sequence linking primary event fields, recorded representations, transformed phantom fields, and resulting event fields. These links are defined by transfer functions whose behavior is restricted by mode-specific and system-imposed limitation sets.

An event is defined as a temporally and spatially localizable occurrence characterized by observable

---

*Address correspondence to this author at the IU International University of Applied Sciences Germany; E-mail: armin.grasnick@iu.org

[1] Perceptual Holistic Analysis and Networked Transfer of Multimodal Adaptive Technologies for Responsive Immersive eXtended Reality.

properties. These properties constitute the only accessible basis for representation. A set of such events forms an event field, which serves as the formal unit of mediation within the framework. The term phantom denotes the data-based representation generated from an event field through recording or synthesis. It is introduced to distinguish this intermediate state from generic notions such as "representation" or "signal." In this work, a phantom field describes a defined, transformable configuration at a specific architectural position between event fields which can be modified independently within the full chain. The term does not denote a perceptual illusion or metaphorical construct, but a technical element of the transfer structure.

Section 2 introduces these components and their mathematical representation. Section 3 defines the overall transfer function of the framework.

2. FORMAL STRUCTURE OF PHANTOMATRIX

An XR system is modeled as a three-stage transfer chain. Physical events are recorded into a phantom data representation. This representation may be computationally modified, and the result is instantiated again as a mediated event field. Each stage operates under specific constraints that determine how event properties are captured, transformed, and presented. The formal notation introduced below expresses these stages and their composition as a mapping between primary and secondary event fields (Figure **1**).

Transformation Architecture

The primary event field E_1 denotes a temporally and spatially distributed set of events characterized by

observable multimodal properties. The recording function r maps the primary event field E_1 to a primary phantom field F_1 under a set of system-specific limitations L:

$$F_1 = r(E_1, L)$$

The phantom field represents the data-based encoding of the recorded event field. The transformation function t operates on F_1 under transformation constraints T and produces a secondary phantom field F_2:

$$F_2 = t(F_1, T)$$

This stage may include algorithmic modification, augmentation, or synthesis.

The display function d maps the secondary phantom field F_2 to a secondary event field E_2 under display-specific limitations D:

$$E_2 = d(F_2, D)$$

Sequential constraint composition defines how event structures emerge from technical mediation. Architectural variation between XR systems depends on how limitations are configured across stages.

Definition of Fields

Primary Event Field E_1

Let $\mathbb{T}$ denote a set of time points and $\mathbb{S}$ a spatial domain. A spatial position is written as $s(x, y, z)$. An event O_i is understood as a temporally and spatially localizable occurrence characterized by observable
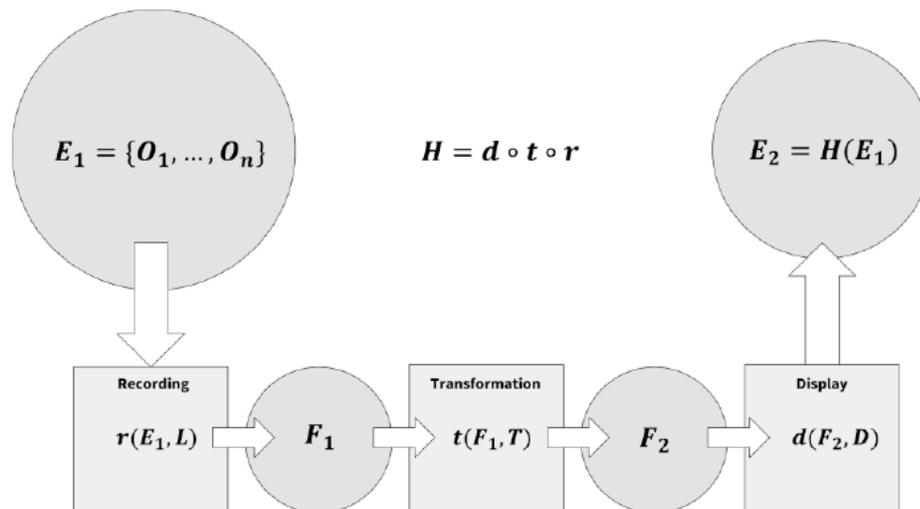

Figure 1: Transformation architecture of the PHANTOMATRIX framework.

Primary event field E_1 is mapped to primary phantom field F_1 by the recording function r under constraints L. The phantom field is modified by the transformation function t under T, resulting in a secondary phantom field F_2. The display function d instantiates the secondary event field E_2 under constraints D. The composed mapping defines the global transfer function $H = d \circ t \circ r$.

properties across different perceptual modes. An event can be written as:

$$O_i(t, s(x,y,z), \{M_1: \{P_{11}, P_{12}, \dots\}, M_2: \{P_{21}, P_{22}, \dots\}, \dots, M_k: \{P_{k1}, P_{k2}, \dots\}\})$$

where

- $i = 1, \dots, n$ indexes events,

- $j = 1, \dots, k$ indexes perceptual modes M_j,

- P_{jm} denotes the m-th recognizable property associated with mode M_j,

- properties may vary with time $t \in \mathbb{T}$ and spatial position $s \in \mathbb{S}$.

This representation states that events are not directly accessible entities; only their observable, mode-dependent properties can be recorded or displayed. The primary event field is defined as a finite set of events

$$E_1 = \{O_i \mid i = 1, \dots, n\}$$

The primary event field therefore formalizes the observable structure of an environment as a set of multimodal events.

Viewing Configuration

Observable properties may depend on the viewing configuration. Let $\mathbb{A}$ denote a set of viewing angles or observation configurations.

For each $\alpha \in \mathbb{A}$, the primary event field can be written as:

$$E_1(\alpha) = \{O_i(\alpha) \mid i = 1, \dots, n\}$$

Each event then becomes:

$$O_i(\alpha) = (t, s(x,y,z), \{ M_1: \{P_{11}(\alpha), \dots \}, \dots, M_k: \{P_{k1}(\alpha), \dots \} \})$$

The parameter α accounts for the fact that observable properties may vary depending on the position or orientation of the observer. The introduction of $E_1(\alpha)$ therefore does not redefine the event field, but expresses that its perceptible properties can change with the viewing configuration.

Primary Phantom Field F_1

The primary event field is not directly accessible to technical systems. Physical events can only be processed once they have been transformed into a data representation. The primary phantom field F_1 formalizes this intermediate representational state. A recording function

$$F_1 = r(E_1, L)$$

maps the primary event field E_1 to a primary phantom field F_1, representing the machine-readable encoding of the captured data. The set L collects recording limitations:

$$L = \{L_M, L_S, L_T, L_Q, L_D\}$$

where:

- L_M denotes mode-specific constraints,

- L_S spatial limitations,

- L_T temporal limitations,

- L_Q signal-quality limitations,

- L_D data-processing limitations.

These limitations restrict which properties of events in E_1 can be recorded and how accurately they are represented in F_1.

Null Configuration

The framework also includes the case of purely synthetic configurations. If no primary event field is present, the input is the empty set:

$$F_1 = r(\emptyset, L)$$

This represents a null input configuration, where no events are recorded and subsequent content must be introduced exclusively by the transformation stage. The inclusion of the empty set ensures that the same formal description can represent both recorded and fully synthetic XR systems. In this formulation, augmented configurations correspond to $E_1 \neq \emptyset$, whereas fully synthetic configurations are represented by $E_1 = \emptyset$. The distinction between AR- and VR-like systems thus arises from the presence or absence of recorded input within the same transformation structure.

Secondary Phantom Field F_2

The primary phantom field represents the technically captured encoding of the primary event field. In many XR configurations, F_1 is not transmitted unchanged. Instead, it may be modified, extended, or partially replaced by computational processes before being presented. This modification is formalized by a transformation function

$$F_2 = t(F_1, T)$$

which maps the primary phantom field F_1 to a secondary phantom field F_2. The set T denotes a collection of transformation classes:

$$T = \{T_{low}, T_{XR}, T_{high}\}.$$

These classes represent distinct categories of transformation:

- T_{low}: low-level enhancement and signal-oriented processing applied to the captured data,

- T_{XR} : content generation or augmentation processes that introduce or replace elements within the data representation,

- T_{high} : higher-level contextual or interaction-dependent transformations that modify the representation according to system state or user-dependent parameters.

The secondary phantom field F_2 represents computationally transformed data derived from F_1. It may contain enhanced, augmented, or entirely synthesized content, depending on the applied transformation set T.

Secondary Event Field

The secondary phantom field F_2 represents a computational data structure resulting from the transformation stage. However, a phantom field remains a data representation and must be rendered by output devices to become observable. This rendering process is formalized by a display function

$$E_2 = d(F_2, D)$$

that maps the secondary phantom F_2 field to a mediated, secondary event field E_2.

The set D collects display-related limitations:

$$D = \{D_M, D_S, D_T, D_Q, D_D\}$$

where:

- D_M denotes mode-specific display constraints,

- D_S spatial limitations,

- D_T temporal limitations,

- D_Q signal-quality limitations,

- D_D data-processing limitations.

These limitations determine how the transformed data are technically instantiated and how the resulting event field differs from the original primary event field. The secondary event field E_2 represents the technically realized configuration derived from the secondary phantom field. It constitutes the event structure made available to an observer through the technical system.

3. GLOBAL TRANSFER FUNCTION

The preceding sections introduced the primary event field E_1 , the recording function r , the transformation function t, and the display function d, together with their associated limitation sets L, T, and D. These components form a sequential processing chain. The overall behavior of the transfer is formalized by the composed mapping

$$H = d \circ t \circ r$$

Accordingly, the secondary event field is given by

$$E_2 = H(E_1)$$

which can be written explicitly in expanded form as

$$E_2 = d(t(r(E_1, L), T), D)$$

The composed transfer function shows that E_2 results from a sequence of constrained operations applied to E_1 . Recording restricts observable properties under L, transformation restructures the encoded data under T , and display determines instantiation under D . The global transfer function describes XR systems as mappings between event fields. Differences between configurations arise from how recording, transformation, and display limitations are parameterized.

Formally, two limiting cases illustrate this structure:

- If $E_1 \neq \emptyset$, the mediated field is derived from recorded input (augmented configurations).

- If $E_1 = \emptyset$, the mediated field results solely from transformation and display processes (fully synthetic configurations).

Both cases are represented within the same composition.

4. DISCUSSION

PHANTOMATRIX is a formal model of XR mediation, not a psychometric UX theory. Experiential effects can be attributed to specific stages of the mediation chain. The examples below illustrate how common UX phenomena can be interpreted within the transfer chain.

Latency Discomfort

Latency-related discomfort in XR is widely associated with motion-to-photon delay and its contribution to cybersickness and perceptual instability. Within PHANTOMATRIX, motion-to-photon delay is represented as the cumulative temporal effect of

sequential limitations introduced by the recording function, the transformation function, and the display function. Each stage contributes its own temporal restrictions, and the overall delay emerges from their composition in the transfer function. From a user experience perspective, discomfort can be interpreted as arising when temporal relations in the mediated event field diverge from the temporal structure implied by user action. The framework allows latency-related effects to be attributed to specific stages of the transfer chain instead of being treated as an undifferentiated system property.

Presence Breaks

Presence can be interpreted as a condition in which the mediated event field preserves a coherent configuration of observable event properties relative to user action [17]. In PHANTOMATRIX terms, this coherence depends on how properties of events in the primary event field are encoded in the phantom field, modified by the transformation function, and instantiated in the secondary event field through the display function. A presence break corresponds to a discontinuity in this mapping. Certain properties remain consistent across stages, while others are selectively reduced, delayed, or reconfigured due to stage-specific limitation sets. The break does not arise from low realism per se, but from asymmetric preservation of event properties within the composed transfer function. By distinguishing whether inconsistencies originate during recording, transformation, or display, the framework localizes presence-related disruptions within the architecture rather than attributing them to the system.

Cognitive Load

Cognitive load in immersive systems has been associated with increased information processing demands and representational complexity in mediated environments [18]. This demand can be related to the density and variability of operations applied to the phantom field. A higher transformation density increases the number of relations that must be tracked to maintain stable interpretation, without necessarily introducing incoherence. Describing XR as a sequence of constrained transformations provides an operational basis for comparing configurations with respect to their potential experiential impact.

Immersion

Immersion has been described as a property of the system rather than a subjective state, reflecting the extent to which a virtual environment replaces or controls sensory input channels [19]. In the transfer process, immersion can be interpreted as the extent to which observable event properties are actively restructured or substituted relative to the recorded input. Rather than equating immersion with subjective involvement, PHANTOMATRIX treats it as a formal characteristic of how comprehensively the transformation stage modifies the phantom field before instantiation in the secondary event field. Extensive substitution, synthesis, or reconfiguration of event properties within the transformation function corresponds to a higher degree of mediation.

5. CONCLUSION

By separating recording, transformation, and display operations, the model makes explicit where constraints arise and how they propagate through the transformation chain. Within this formulation, augmented configurations correspond to systems in which $E_1 \neq \emptyset$, whereas fully synthetic configurations are represented by $E_1 = \emptyset$. Hybrid systems emerge as intermediate cases that differ only in input conditions and transformation parameters. AR- and VR-like systems are thus expressed within a single compositional structure.

By distinguishing between physical events, data representations, and technically instantiated event fields, PHANTOMATRIX provides a formal vocabulary for describing XR architectures independently of specific hardware implementations. This presentation enables systematic comparison of XR configurations across architectural stages. Beyond descriptive consistency, the framework enables analysis of XR systems. Constraint sets can be examined to identify stage-specific limitations, transformation classes can be analyzed with respect to their influence on mediated event fields, and alternative configurations can be expressed within the same mapping principle. In this way, our approach establishes a formal basis for comparative analysis and systematic modeling of XR architectures. Future work may refine the mathematical characterization of limitation sets, develop quantitative measures for transformation stages, or apply the framework to evaluations of existing XR systems.

From a design perspective, stage-specific attribution of constraints allows XR systems to be analyzed with respect to their architectural impact on experience. Perceptible inconsistencies or variations in mediated environments can be associated with recording constraints, transformation strategies, or display limitations rather than being treated as undifferentiated system effects. By mapping design decisions to distinct components of the global transfer function, UX-related modifications can be conceptualized as parameter changes within recording, transformation, or display stages.

REFERENCES

[1] Clua E, Porcino T, Trevisan D, Cardoso JCS, Lisboa T, Peres V, *et al*. Workshop: Challenges for XR in Digital Entertainment. In: Baalsrud Hauge J, C. S. Cardoso J, Roque L, Gonzalez-Calero PA, editors. Entertainment Computing - ICEC 2021 [Internet]. Cham: Springer International Publishing; 2021 [cited 2023 May 6]. p. 489-98. https://doi.org/10.1007/978-3-030-89394-1_43

[2] Jahn K, Oschinsky FM, Kordyaka B, Machulska A, Eiler TJ, Gruenewald A, *et al*. Design elements in immersive virtual reality: the impact of object presence on health-related outcomes. INTR. 2022; 32: 376-401. https://doi.org/10.1108/INTR-12-2020-0712

[3] Plechatá A, Makransky G, Böhm R. Can extended reality in the metaverse revolutionise health communication? npj Digit Med. 2022; 5: 132. https://doi.org/10.1038/s41746-022-00682-x

[4] Vasarainen M, Paavola S, Vetoshkina L. A Systematic Literature Review on Extended Reality: Virtual, Augmented and Mixed Reality in Working Life. IJVR. 2021; 21: 1-28. https://doi.org/10.20870/IJVR.2021.21.2.4620

[5] Billinghurst M, Duenser A. Augmented Reality in the Classroom. Computer. 2012; 45: 56-63. https://doi.org/10.1109/MC.2012.111

[6] Ziker C, Truman B, Dodds H. Cross Reality (XR): Challenges and Opportunities Across the Spectrum. In: Ryoo J, Winkelmann K, editors. Innovative Learning Environments in STEM Higher Education [Internet]. Cham: Springer International Publishing; 2021 [cited 2023 May 6]. p. 55-77. https://doi.org/10.1007/978-3-030-58948-6_4

[7] Moses R, Garia N, Devan P. Digital reality - A technical primer. Deloitte Insights; 2018.

[8] Felnhofer A, Kothgassner OD, Schmidt M, Heinzle A-K, Beutl L, Hlavacs H, *et al*. Is virtual reality emotionally arousing? Investigating five emotion inducing virtual park scenarios. International Journal of Human-Computer Studies. 2015; 82: 48-56. https://doi.org/10.1016/j.ijhcs.2015.05.004

[9] Meuleman B, Rudrauf D. Induction and Profiling of Strong Multi-Componential Emotions in Virtual Reality. IEEE Trans Affective Comput. 2021; 12: 189-202. https://doi.org/10.1109/TAFFC.2018.2864730

[10] Somarathna R, Bednarz T, Mohammadi G. Virtual Reality for Emotion Elicitation - A Review. IEEE Trans Affective Comput. 2022; 1-21.

[11] Al-Jundi HA, Tanbour EY. A framework for fidelity evaluation of immersive virtual reality systems. Virtual Reality. 2022; 26: 1103-22. https://doi.org/10.1007/s10055-021-00618-y

[12] Skarbez R, Smith M, Whitton MC. Revisiting Milgram and Kishino's Reality-Virtuality Continuum. Front Virtual Real. 2021; 2: 647997. https://doi.org/10.3389/frvir.2021.647997

[13] Stauffert J-P, Niebling F, Latoschik ME. Latency and Cybersickness: Impact, Causes, and Measures. A Review. Front Virtual Real. 2020; 1: 582204. https://doi.org/10.3389/frvir.2020.582204

[14] Warburton M, Mon-Williams M, Mushtaq F, Morehead JR. Measuring motion-to-photon latency for sensorimotor experiments with virtual reality systems. Behav Res [Internet]. 2022 [cited 2023 May 13]; https://doi.org/10.1101/2022.06.24.497509

[15] Zhao N, Zhang X, Noah JA, Tiede M, Hirsch J. Separable Processes for Live "In-Person" and Live "Zoom-like" Faces. Imaging Neuroscience [Internet]. 2023 [cited 2023 Nov 4]; https://doi.org/10.1162/imag_a_00027

[16] Grasnick A. Transfer Functions and Event Fields in XR: The PHANTOMATRIX Framework for Quantifying Perception Convergence (Preprint) [Internet]. 2023 [cited 2024 June 5]. https://doi.org/10.21203/rs.3.rs-3556886/v1

[17] Slater M. Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. Phil Trans R Soc B. 2009; 364: 3549-57. https://doi.org/10.1098/rstb.2009.0138

[18] Makransky G, Petersen GB. The Cognitive Affective Model of Immersive Learning (CAMIL): a Theoretical Research-Based Model of Learning in Immersive Virtual Reality. Educ Psychol Rev. 2021; 33: 937-58. https://doi.org/10.1007/s10648-020-09586-2

[19] Slater M. Immersion and the illusion of presence in virtual reality. British J of Psychology. 2018; 109: 431-3. https://doi.org/10.1111/bjop.12305