An Ontological Conceptual Model for Structuring Multimodal User Behavior in Virtual Reality*

Ítalo Oliveira[1,2,*], Delaram Javdani Rikhtehgar[3] and Shenghui Wang[3]

[1]Semantics, Cybersecurity, and Services (SCS), University of Twente, Drienerlolaan 5, 7522 NB, Enschede, The Netherlands
[2]Y.digital, Arnhemse Bovenweg 140, 3708 AH, Zeist, The Netherlands
[3]Human Media Interaction (HMI), University of Twente, Drienerlolaan 5, 7522 NB, Enschede, The Netherlands

Abstract

Virtual Reality (VR) is increasingly employed in applications such as virtual exhibitions, safety training, and medical simulations. Understanding user behavior in these environments is essential for enhancing user experience and informing interaction design. However, analyzing VR behavior is complex due to the multimodal nature of interactions— encompassing gestures, gaze, and speech— which results in vast and heterogeneous data streams. Effectively managing and interpreting this data remains a significant challenge. Well-founded ontologies offer a structured approach to understanding, organizing, and analyzing behavioral data, yet their application in VR behavioral research remains underexplored. To address this gap, we introduce *OnBehaVR*— an ontology designed to represent and analyze user behavior in VR. Developed using the Unified Foundational Ontology (UFO) and the OntoUML conceptual modeling language, OnBehaVR provides a formal framework for conceptualizing VR interactions. It aims to (1) clarify the conceptualization of user behavior in VR, (2) enable data integration within this domain, and (3) facilitate behavior analysis through automated reasoning and semantic queries. By leveraging ontology-driven approaches, OnBehaVR contributes to the systematic study of multimodal interactions and supports the development of adaptive and personalized virtual experiences.

Keywords

Virtual Reality, User Behavior, Human-Computer Interaction, OntoUML, Unified Foundational Ontology

1. Introduction

Virtual Reality (VR) environments offer users immersive and interactive experiences and are increasingly applied in domains such as virtual exhibitions [1], safety training [2], and medical simulations [3]. Understanding user behavior within these environments is crucial for improving user experience and interaction design. However, gaining this understanding is challenging due to several factors: the complex interplay between physical and virtual elements; the multimodal nature of user behavior in VR, which involves gestures, gaze, speech, and other heterogeneous data streams from various sensors; and the diversity of VR platforms and use cases, which complicates data interpretation and integration. In this context, a central research question arises: *How can we effectively interpret and manage data to better understand user behavior in VR?*

A comprehensive understanding of user behavior in VR is a matter of *ontology*, as it requires an application-independent representation of the domain [4]. This conceptual insight can be used to structure and analyze user behavior data in VR through various means, such as Semantic Web technologies, and across different VR systems. However, this approach to user behavior in VR has yet to be presented. To fill this gap, we introduce OnBehaVR— a well-founded ontology designed to represent and analyze user behavior in VR. Developed using the Unified Foundational Ontology (UFO) and the OntoUML

ER2025: Companion Proceedings of the 44th International Conference on Conceptual Modeling: Industrial Track, ER Forum, 8th SCME, Doctoral Consortium, Tutorials, Project Exhibitions, Posters and Demos, October 20-23, 2025, Poitiers, France
*Corresponding author.
i.j.dasilvaoliveira@utwente.nl (Í. Oliveira); d.javdanirikhtehgar@utwente.nl (D. J. Rikhtehgar); shenghui.wang@utwente.nl (S. Wang)
https://italojsoliveira.github.io/ (Í. Oliveira); https://people.utwente.nl/d.javdanirikhtehgar (D. J. Rikhtehgar); https://people.utwente.nl/shenghui.wang (S. Wang)
0000-0002-2384-3081 (Í. Oliveira); 0009-0007-8266-8395 (D. J. Rikhtehgar); 0000-0003-0583-6969 (S. Wang)

conceptual modeling language, OnBehaVR provides a formal framework for conceptualizing VR interactions. Reusing gUFO [5], OnBehaVR has also been implemented in the Web Ontology Language (OWL) to support semantic web applications. OnBehaVR primarily aims to: (1) provide conceptual clarification of user behavior in VR; (2) enable data integration within this domain; and (3) facilitate user behavior analysis through automated reasoning and semantic queries.

To illustrate its applicability, we apply OnBehaVR in a virtual exhibition case study (as in [6]), using eye-tracking data to examine how ontology-based modeling can help structure and integrate behavioral data. This study offers initial insights into the potential of ontologies to enhance the semantic organization of user interactions in VR environments, providing a foundation for further investigation into engagement patterns in virtual spaces. By leveraging ontology-driven approaches, OnBehaVR supports the systematic study of multimodal interactions and lays the groundwork for advancements in adaptive and personalized VR experiences.

The paper is structured as follows: Section 2 discusses related works concerning VR user behavior ontologies. Section 3 presents the motivation scenario (Section 3.1) and the explicit requirements (Section 3.2) behind our ontology. Section 4 introduces our contribution after methodological considerations (Section 4.1): OnBehaVR, an ontology for structuring user behavior in VR environments. OnBehaVR is divided into two views: (a) VR user behavior experience, described in Section 4.2, and (b) VR user record, detailed in Section 4.3. Section 5 evaluates how OnBehaVR satisfies the requirements proposed in Section 3.2. Finally, Section 6 presents our concluding remarks and directions for future work.

2. Related Work

Few research works develop ontologies specifically to address VR user behavior— a subset of human-computer interaction studies. A 2021 systematic literature review on ontologies in human-computer interaction by Costa et al. [7] identified 35 ontologies, but only one [8] pertains to virtual reality. The review observed that, overall, these ontologies do not follow good practices for ontology engineering, as outlined by D'Aquin and Gangemi [9], namely: (i) reusing foundational ontologies; (ii) being formally rigorous; (iii) implementing non-taxonomic relations (in addition to taxonomic ones); (iv) being modular; (v) implementing an international standard; (vi) being based on competency questions; and (vii) following an evaluation method [7]. The aforementioned VR ontology [8] satisfies only (iv) modularity and (vii) compliance with an evaluation method. We have observed that these deficiencies are common among other VR ontologies not included in this literature review. In general, they do not adhere to the FAIR principles [10]. In follow-up work, Costa et al. [11] introduced a Human-Computer Interaction Ontology grounded in the Unified Foundational Ontology (UFO). It addresses the gaps identified earlier and is proposed as a reference model to facilitate understanding and communication within the domain. However, it does not address our concerns, which are more specific to the VR environment— particularly regarding how to interpret and structure multimodal data.

Messaoud et al. [12] present a proposal that resembles ours, as they build ontologies to semantically structure a virtual environment, along with the behaviors and interactions of entities. The goal of their semantic layer is to ease the design of intelligent virtual environments and agents. This intelligence involves the capacity of artificial agents to exhibit human-like behaviors and to assist users in solving specific problems. Such behavior is executed via semantic queries and guidelines. The ontologies include notions related to the virtual environment— such as objects, geometry information, and position— as well as virtual human concepts, including height, age, gender, and personality. Although the authors claim that their "architecture has been successfully tested in 3D dynamic environments," no demonstration is provided, and the OWL ontologies are unavailable. These appear to have been built from scratch without foundational references or reuse of existing ontologies.

Pacheco et al. [13] propose two ontologies to describe 3D objects in a static VR environment, based on concepts from VRML and Java 3D— formal languages used to describe 3D objects. However, the ontologies are neither described nor publicly available for proper assessment. Even the ontology representation language is unclear. Cha et al. [14] build an OWL ontology to represent Google Art

Project metadata about artworks. Their goal is to provide useful information to learners in a VR environment. Like Pacheco et al. [13], they do not follow established best practices for ontology engineering [9], nor do they provide implementation details.

Sokołowski and Walczak [8] use OWL to create one ontology describing the interaction channels available in a particular VR system configuration, and another ontology describing the interaction capabilities within the VR environment. These ontologies support configuration mappings that facilitate the migration of an application from one environment to another. For example, a “Cancel” action in a VR environment can be mapped to the “X” button on a game controller. The authors address a different problem from ours. The choice of OWL is not justified; they appear to use it primarily for its taxonomic structure, and the ontologies are neither described nor made available.

In the context of the *Plausible Representation of Emergency Scenarios for Training Operations* (PRESTO) research project, Dragoni et al. [15] develop OWL ontologies to describe artificial entities and their behaviors in gaming, intending to decouple the description of VR scenarios from their physical implementation, which is handled by developers. The ontologies are built upon the *Descriptive Ontology for Linguistic and Cognitive Engineering* (DOLCE), extended with domain-specific concepts learned from VR experts and environments (XVR). The ontology design is guided by questions such as “What are the entities that exist, or can be said to exist, in a Virtual Reality scenario?” and “How can such entities be grouped, related within a hierarchy, and subdivided according to similarities and differences?” By abstracting VR tool-specific details, the PRESTO ontology helps simplify VR scenario development. However, due to copyright constraints, the ontology cannot be published, which limits our ability to evaluate it.

Chokwitthaya et al. [16] present the *Virtual Human-Building Interaction Experimentation Ontology* (VHBIEO) to support the standardization and replicability of virtual human-building interaction experiments. *Human-Building Interaction* is an interdisciplinary field that studies the dynamics between humans and the built environment, aiming to enhance comfort, usability, and energy efficiency. VHBIEO reuses concepts from several ontologies, including the *Ontology of Scientific Experiments* (EXPO), the *Suggested Upper Merged Ontology* (SUMO), the *Semantic Sensor Network Ontology* (SSN), and others, and follows the DOGMA ontology engineering methodology. Competency questions were defined to elicit VHBIEO’s requirements, including: “What are the human-building interactions involved in an experiment?”, “What are the building parameters involved in an experiment?”, “Who performs interactions?”, and “What are the devices used to observe and collect data?”, among others. VHBIEO was built using the Protégé editor and is publicly available[1]. This ontology is by far the most comprehensive and best designed, adhering to good practices in ontology engineering and the FAIR principles, and stands as an exception among the VR ontologies reviewed. However, it is a domain ontology specifically focused on virtual human-building interaction experimentation, which is not our use case.

3. Ontology Requirements

Following the design science methodology [17], we aim to (1) produce an information artifact (OnBehaVR), (2) in order to address an important research problem (understanding and modeling user behavior in VR), (3) following rigorous methods of artifact construction (ontological foundations); (4) we explicitly define requirements for the artifact evaluation, (5) and will show that the artifact satisfies them; (6) finally, we communicate our results via documentation.

3.1. Motivation

Consider the following motivational scenario (one among many possible related scenarios), clarifying which individuals are relevant in our domain of interest. In a VR user behavior research experiment or during a casual virtual exhibition visit, a user wears a VR headset connected to a computer running a VR application built on a game engine (Unity, Unreal, Godot, Open 3D, etc.). This VR connection

[1] https://w3id.org/vhbieo.

initiates an immersive experience in which the user may create an account and at least one avatar and enter personal data (e.g., name, sex, language). The user likely has some goals in mind, even if it is merely curiosity or the desire to find something interesting. In less casual scenarios, such as factory safety training, the user may intend to learn specific subjects or skills. The user experience, composed of various events, may or may not satisfy those goals to some degree. This experience involves interactions with objects and the environment, including paintings, doors, rooms, and NPCs, such as a virtual museum guide. The user's behavior reflects their mental state during each interaction. For example, if the user is interested in a particular art piece, they may focus their gaze on it and ask the guide questions about it. While the experience is running, devices capture multimodal raw data, including conversations and gaze tracked by the headset, heartbeats monitored by a smartwatch, and cerebral activity measured by electroencephalography (EEG) sensors.

Given this scenario, researchers, designers, and developers are interested in understanding user behavior in the VR environment: What virtual objects did a user's avatar interact with during a VR session? What are the users' feelings and mental states during different interactions? Why did the user behave that way? Was the user's goal satisfied by the VR experience? What can be inferred from all the collected data?

3.2. Requirements

To answer questions like these, we need an application-independent *ontological conceptual model* for structuring, interpreting, and integrating multimodal user behavior in VR. With this in mind, we define both functional and quality requirements for OnBehaVR as an information artifact. The former establishes the scope of OnBehaVR by defining what it is expected to do. The primary functional requirement of OnBehaVR is to offer *conceptual clarification* about the domain. This elucidation shall support two other tasks:

1. *Ontology-based data integration*: OnBehaVR shall structure information about users' behavior in VR environments so that data from different sources can be explained, represented, and integrated. Following Guizzardi and Guarino [18], the notion of explanation here means that OnBehaVR shall make explicit the content of human communications, systematically revealing the ontological commitment (i.e., truthmakers) underlying symbolic descriptions of our domain of interest (user behavior in the VR environment). Thanks to this explanatory power, data representation and integration become possible. Data representation means that implementations of OnBehaVR can store data points related to user behavior in VR. Data integration means that specific implementations of OnBehaVR can aggregate data from different sources under a unified theory of what the data is about. Such data may include eye-gaze tracking, virtual object locations, segmentation, areas of interest, object information, user demographic data, interactions with virtual agents (e.g., choices, dialogues), and information from smart devices (e.g., heart rate), among others.
2. *Knowledge-based systems for intelligent agents*: OnBehaVR shall support knowledge representation and automated reasoning about users' behavior in VR environments. The idea, similar to that in [12], is that knowledge-based systems can be developed using this support to enable virtual intelligent agents that interact meaningfully with users.

The ontology engineering literature discusses many quality criteria for ontologies [19]. These criteria are desiderata to be pursued in ontology engineering, as they prescribe attributes that enable or enhance the ontology's functionalities. Their selection and interpretation may vary across different authors. Aligned with the literature, we define the following quality criteria:

1. *Domain adequacy*: OnBehaVR shall accurately represent shared, real-world conceptual elements of user behavior in VR environments [4].
2. *Documentation*: OnBehaVR shall include explicit definitions for all its concepts and be accompanied by publicly available documentation.

3. *Consistency*: OnBehaVR shall be free from logical contradictions.
4. *Ontological coherence*: OnBehaVR shall adhere to upper-level ontological distinctions.
5. *Extensibility*: OnBehaVR shall support extensions to integrate data from new sources when needed.
6. *FAIR principles* [10]: OnBehaVR shall be findable, accessible, interoperable, and reusable, following the guidelines of the International Conference on Formal Ontology in Information Systems for research artifacts [20].

4. OnBehaVR: An Ontology for Structuring User Behavior in Virtual Reality

OnBehaVR captures certain philosophical assumptions about VR. According to David Chalmers [21], *a virtual reality environment is an immersive, interactive, computer-generated environment.* An immersive environment generates a perceptual experience from a perspective within it, usually through a three-dimensional visual experience. An environment is interactive when user actions make a practical difference to what happens within it. An environment is computer-generated if it is grounded in a computational process, such as a computer simulation. "Virtual reality" as a count noun is synonymous with "virtual reality environment," while as a mass noun, it encompasses both virtual reality environments and virtual reality technology. We agree with Chalmers that virtual objects are digital objects grounded in computational processes— processes that are themselves grounded in physical operations on one or more computers. In this sense, virtual objects are physical objects. However, from an information modeling perspective, they possess different ontological natures. We maintain that virtual objects are disjointed from physical objects because, for instance, a virtual chair contrasts with a physical chair. According to Brey [22], *a virtual object is a digital object that is represented graphically as an object or region in a 2D or 3D space and that can be interacted with or used through a computer interface.* That said, virtual objects constitute a reality realm among all possible objects, partially mirroring physical and social objects. Virtual events are events wherein virtual objects participate, occurring within a VR environment.

4.1. Methodology and Design Choices

To build *OnBehaVR*, an ontological model of user behavior in VR environments, we proceeded as follows: (1) we selected UFO [23], a well-known upper ontology, as the basis for our domain model; (2) we employed OntoUML[2], a UFO-based, ontology-driven conceptual modeling language, to describe our conceptual model; (3) we took into account (a) the philosophical literature on virtual reality, (b) similar models (mentioned in Section 2), (c) the dataset schemata relevant to our data integration problem, and (d) expert knowledge of human-computer interaction in VR environments; and (4) we built OnBehaVR to satisfy the explicit functional and quality requirements defined earlier. We selected UFO as our upper ontology because it has been widely used in socio-technical domains [24], and it has its own conceptual modeling language (OntoUML) with several built-in services (such as syntax checking, OWL generation, JSON serialization, and more) [25]. The ontological distinctions of UFO are embedded in OntoUML through the UML profile mechanism applied to the class diagram. As a result, OntoUML models are constructed through the iterative instantiation of ontology design patterns, each representing a UFO micro-theory. In this way, OntoUML models conform to UFO's ontological partitioning, which is designed to prevent certain modeling errors— such as mixing events, objects, and intrinsic properties.

OnBehaVR is publicly available in multiple formats (figures in `.jpeg`, and editable files in `.vpp`, `.json`, `.ttl`, and `.html` for different applications) at: https://w3id.org/virtual-reality-ontology/git.

[2]OntoUML diagrams are supported through a plugin for Visual Paradigm. See: https://github.com/OntoUML/ontouml-vp-plugin.

4.2. Virtual Reality User Experience

The OnBehaVR's taxonomy begins with a distinction between Physical Objects and Virtual Objects[3]. The latter can represent other objects, including additional Virtual Objects; although, typically, Virtual Objects represent Physical Objects. For example, a virtual painting may represent an actual (physical) painting. However, Virtual Objects may also exist without representing any "exterior" thing— for instance, a virtual art piece created specifically for a virtual exhibition. Furthermore, Virtual Objects can represent social objects such as money, contracts, and marriage. Together, social objects, Physical Objects, and Virtual Objects are generalized by a class termed Generic Object— an abstract class used to denote UFO Substantials.

An Agent is a notion from UFO, meaning an object that bears certain Intentional Moments— namely, Beliefs, Desires, and Intentions [26]. A Person is a subtype of Agent that is also a Physical Object. This implies that our definition of a Person excludes fictional individuals. Nevertheless, entities such as companies, robots, and virtual non-player characters (NPCs) can be regarded as Agents, provided it is meaningful to attribute to them Beliefs, Desires, or Intentions. Thus, a Virtual Agent is an Agent that is also a Virtual Object. For example, multiple NPCs that are responsible for different things— such as a guard protecting a door, a nurse that can cure the user's Avatar, and a virtual guide accompanying the user through the environment— can all be considered Virtual Agents. In contrast, virtual paintings and virtual doors are not agentic in this sense.

Virtual Dispositions are dispositions (intrinsic properties) of Virtual Objects, manifested through Virtual Events. Virtual Dispositions determine what a Virtual Object can do within a VR environment. For example, a virtual door may possess the Virtual Disposition to open when triggered by a Virtual Agent's actions. Additionally, Virtual Dispositions define the possible behaviors of an Avatar, a graphical representation of the user. An Avatar is a Virtual Agent standing for a Person who plays the role of a Virtual Reality User. Virtual Dispositions are important because redesigning the VR environment frequently means redefining Virtual Dispositions, i.e., how objects are expected to behave in certain conditions.

A Virtual Actor is a Virtual Agent defined by its participation in a Virtual Event (Virtual Action). Often, in virtual games, the user's Avatar assumes certain roles or receives recognition due to completing specific tasks or quests. These roles or recognitions are subtypes of Virtual Actor. A Virtual Reality User is defined as a Person who participated in a VR User Experience and is related to a Virtual Reality Connection. For this reason, a Virtual Reality User is considered a historical role. As a Person, the Virtual Reality User has Mental states, psychological intrinsic properties. These include distraction, boredom, confusion, interest, attention, curiosity, sadness, and many others [27]. In the context of user behavior research, certain mental states can be inferred from users' physical and virtual behaviors. For example, a prolonged gaze on a painting may suggest the Virtual Reality User is interested in the artwork. An accelerated heart rate may indicate excitement. For this reason, VR User Experiences are considered manifestations of Mental states.

A Virtual Reality Program is the technological means that provides access to a VR environment. Therefore, a Virtual Reality Connection must be associated with a Virtual Reality Program— for example, a VR environment (such as a virtual museum) created using the a game engine. A Virtual Reality Connection aggregates intrinsic properties by mediating specific components, namely: a Virtual Reality Program, Devices, a Virtual Reality User, Virtual Objects, and an Avatar. The manifestation of these properties (including Mental states) results in VR User Experience events. For instance, when a person logs in to a VR environment to control an Avatar in a virtual art exhibition, they establish a Virtual Reality Connection.

A Device is any piece of hardware that provides data to a VR system, such as heart rate monitors or eye-gaze trackers. An Immersive Device generates a three-dimensional image that appears to surround the user. The presence of such a device is essential to defining a VR User Experience. VR headsets are examples of Immersive Devices.

[3]In this paper, expressions in small caps font, such as Physical Objects and Virtual Objects, refer to explicit classes within OnBehaVR.

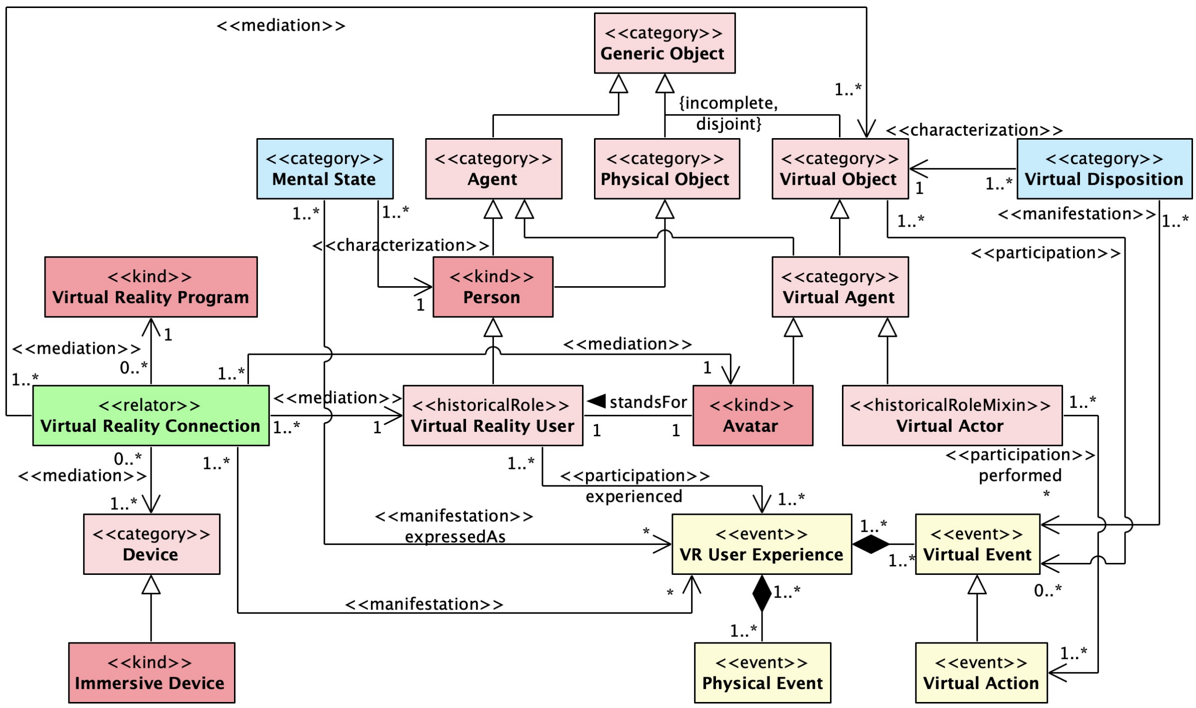

Figure 1: Virtual Reality User and Experience. The OntoUML stereotypes connect types and relations in these models to ontological categories of monadic and relational universals in UFO, respectively [23]. The colors in these diagrams represent a convention used by the OntoUML community: object types are represented in pinkish, intrinsic aspect types in blue, and event types in yellow, and truth-makers of material relations are represented in green.

Physical Events include a Person's physical movements and heartbeats. They are disjoint from Virtual Events but together constitute the VR User Experience. Virtual Events are manifestations of Virtual Objects' dispositions. They are occurrences within a VR environment. In UFO, all events are manifestations of dispositions— intrinsic properties inhering in objects. The participation of objects in events *derives* from these manifestations. A Virtual Action is an event performed by a Virtual Agent *because of* their Intentional Moments (Intentions, Beliefs, and Desire). To clarify these notions, consider the following example: a virtual door, a Virtual Object, manifests its Virtual Dispositions by opening (Virtual Event); however, an Avatar, because it *intends* to do so, performs a Virtual Action by opening the virtual door. The opening is a complex event composed of the manifestations of dispositions inhering in both the door and the Avatar. A VR User Experience is a complex event composed of Physical Events and Virtual Events. The multimodal user interaction comes from this complex composition of the VR User Experience.

It is important to realize that the Virtual Reality User as a Physical Object can only produce Physical Events. However, the Avatar, controlled by the Virtual Reality User, is a Virtual Object that can produce Virtual Events. The dynamic interplay between Physical Objects and Virtual Objects participating in different events explains what the *interaction* in a VR environment is, forming a VR User Experience. Figure 1 summarizes these ideas in the main view of OnBehaVR.

4.3. Virtual Reality User Record

OnBehaVR distinguishes between, on the one hand, the VR User Experience as a complex event composed of Physical and Virtual Events, and, on the other hand, the VR User Record associated with these events. These records, registered by different Devices (VR headsets, smartwatches, etc.), correspond to multimodal data about the Virtual Reality User, including eye-gaze data, heartbeat rate, dialogs, etc. Abstract Devices are implemented algorithms or procedures that generate Derived VR User Record, which is a VR User Record calculated from certain Record Feature values. A

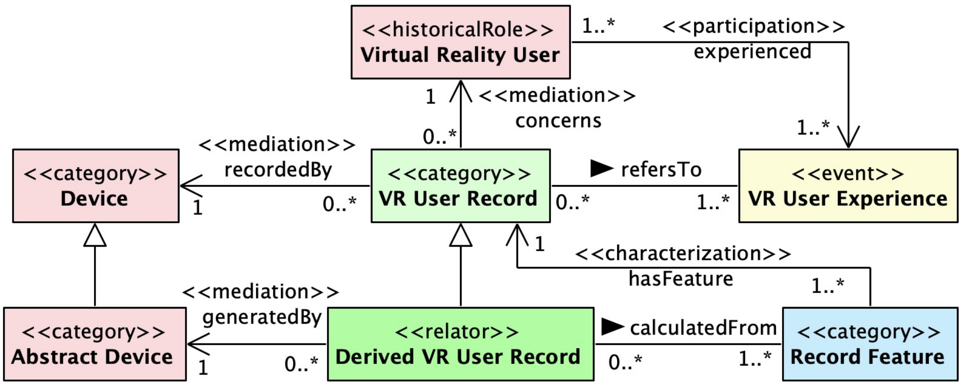

Figure 2: Multimodal Virtual Reality User Record.

Record Feature denotes any quality of interest about the VR User Record. For example, the geometry of an area of interest whose values are coordinates. Record Features can also be understood and represented as *attributes* of VR User Records. Eye-gaze fixation duration is an example of a Derived VR User Record made out of eye-gaze coordinates and timestamps.

In other words, one thing is an event, and another is a recorded measurement related to this event. The utility of this conceptualization is that we can *group* record series as events for analysis. For example, given a series of records with timestamps between 1 second and 40 seconds, we may notice this period regards the event wherein the user's Avatar visited a specific virtual room. Cluster algorithms can also help build events like this and discover insights. Figure 2 depicts the OnBehaVR view of the VR User Record.

The main purpose of this view is to structure data from different sources and integrate them into the ontological theory depicted in Figure 1. Our model intends to adapt and generalize insights from the snow measurement representation in the AlpineBits DestinationData[4]. It is purposely general to achieve the necessary flexibility to represent multimodal data. To do that, we need to create subtypes of the VR User Record corresponding to the type of data we are interested in. To understand how this can be executed, consider several common user behavior records associated with eye-gaze (see [28] for a detailed description) in the context of OnBehaVR:

- **Demographics**: These are attributes of Person and, consequently, Virtual Reality User, displayed in Figure 1. It includes name, gender, age, nationality, native language, and others. These data are not VR User Records, but the latter concern Virtual Reality User with those attributes.
- **VR-generated records**: They are based on the headset's capabilities, usually exportable in `.csv` or similar format, including:
  - *Timestamp*: It starts counting in milliseconds from the establishment of a Virtual Reality Connection by a Virtual Reality User. Timestampt can be converted to date format if needed. It is an attribute of the respective record subtype.
  - *Player position*: The player position in a VR environment is defined by 3D coordinates. It is a subtype of VR User Record.
  - *Gaze position*: It concerns eye-gaze fixation described by 3D coordinates. It is a subtype of VR User Record.
  - *Object gaze*: It is defined by 2D Virtual Object coordinates, considering the Virtual Object between [0,0] top left and [1,1] bottom right. It is a subtype of VR User Record.
- **Derived VR User Records**:

[4]This standard concerns data exchange for tourist destinations such as events and ski areas in the Alps. See: https://www.alpinebits.org/open-standard/destination-data/.

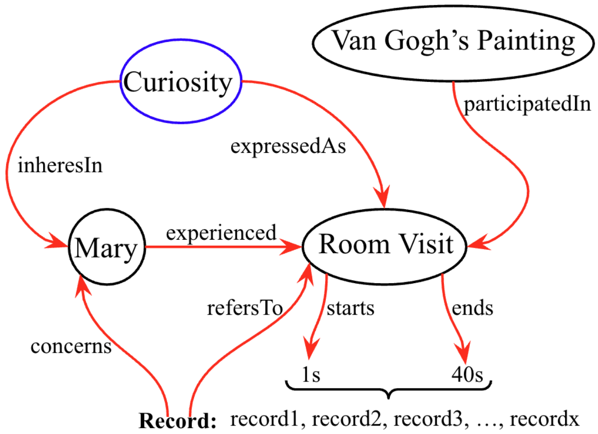

Figure 3: An illustration of a user behavior understanding in VR based on OnBehaVR.

- *Transition*: The frequency of eye movement transitions between Areas of Interest in a Virtual Object (e.g., a virtual painting).
- *Visited area of interest*: The count of Areas of Interest attended.
- *Fixation duration*: The duration of an eye-gaze fixation, measured in milliseconds. It is an attribute of eye-gaze fixation. It is defined by the difference between timestamps.

As a conceptual model, OnBehaVR does not consider implementation details, such as the datatypes of records and attributes. In fact, there are multiple useful ways to implement OnBehaVR by aligning it with OWL ontologies related to measurements, such as the Ontology of units of Measure (OM) [29][5] and Semantic Sensor Network Ontology (SSN)[6] to specify VR User Record, and GeoJSON[7] for representing geospatial data. OnBehaVR offers conceptual insights about how to structure, integrate, and analyze user behavior multimodal data in VR. A full implementation is outside the scope of this paper since it involves a different open problem— the ontology alignment between gUFO [5], SSN, OM, and geospatial ontologies.

To better visualize this, consider Figure 3 illustrating a user behavior *understanding* in VR based on OnBehaVR. A `Room Visit` is a Virtual Event composing `Mary`'s VR User Experience wherein a `Van Gogh's Painting` also participates. The `Room Visit` started at 1 second and ended at 40 seconds after the initiation of the Virtual Reality Connection. During this time frame, many VR User Records were registered concerning `Mary`: `record1, record2, record3,..., recordx`. For the sake of argument, suppose that studies have shown that specific *patterns*[8] present in these records in such context reflect the user's `Curiosity`— a `Mary`'s Mental State. This understanding can inform the redesign of the VR environment to stimulate the user's curiosity and personalize the experience. For example, in virtual museum contexts, conversational agents have been integrated to serve as knowledgeable guides to enhance the accessibility of information [6]. Such Virtual Agents can adaptively respond to users according to that understanding.

[5]https://www.wur.nl/en/product/ontology-of-units-of-measure-om.htm.
[6]https://www.w3.org/TR/vocab-ssn.
[7]GeoJSON is a geospatial data interchange format based on JavaScript Object Notation (JSON). See: https://datatracker.ietf.org/doc/html/rfc7946.
[8]Discovering these patterns is an open empirical question for human-computer interaction research.

5. Evaluation

Considering the requirements defined in Section 3.2[9], we argue that OnBehaVR satisfies them to a certain degree. Let us consider the conceptual clarification power of OnBehaVR. As the first well-founded ontological conceptual model for user behavior in VR, OnBehaVR brings insights and structure to this area of the Human-Computer Interaction field. Several concepts that are often informally adopted have been articulated in a real-world semantics. For example, the *interactivity* of the VR environment emerges from the succession of Physical and Virtual Events composing the VR User Experience. These events are manifestations of multiple interrelated capabilities (dispositions) of objects of different kinds (virtual and physical objects). Events may impact positively or negatively Agent's Intentions, a key question to be answered when (re)designing VR environments. In this context, a natural query is "What virtual objects did a user's avatar interact with during a VR session?". This query can be formulated in OnBehaVR by searching for the objects that have participated in the events of that VR session, according to the records. And this participation is derived from the manifestation of those objects' capabilities.

The requirement of data integration support is fulfilled because we can map multimodal data to relevant concepts of OnBehaVR. This enables data integration via RDF knowledge graph construction, gathering datasets' information. The OnBehaVR model is general enough to be extended to capture data from new sources, provided that the VR User Record is correctly specialized (for example, by creating subclasses of the VR User Record for Player Position and Transition). More importantly, OnBehaVR *explains* the relations between recorded user data, user experience, user mental states, and several other domain-specific elements.

The support for knowledge-based systems for intelligent agents is achieved for two reasons: (a) OnBehaVR is a theory describing users' behaviors in VR, therefore it constitutes explicit domain knowledge that can be embedded into a knowledge-based system; (b) OnBehaVR already has a computational language representation in the Web Ontology Language (OWL), generated by a service of OntoUML and serialized in the Turtle format. In other words, OnBehaVR satisfies this requirement both content-wise and computation-wise. Both capabilities can be combined to inform the virtual intelligent agents in their interaction with the user's Avatar in a real-world system.

Regarding the quality requirements (QR), OnBehaVR satisfies domain adequacy (QR1) thanks to the methodology of its construction, relying on a well-founded upper ontology and modeling language, philosophical expertise in VR, domain datasets, and iterations with domain experts. Every concept of OnBehaVR has been explicitly defined, and each description is available across multiple formats of our ontology. This satisfies QR2 concerning good documentation. Because OnBehaVR has an OWL version, we have proven it is logically consistent and satisfiable (QR3) by automated reasoning[10]. OnBehaVR complies with upper ontological distinctions (QR4) because it is based on UFO and OntoUML, as we already explained. In fact, this is guaranteed by the OntoUML syntax checking service. OnBehaVR satisfies QR5 (extensibility) because it can be tailored for multiple VR domains, as well as for capturing data from new sources. Finally, OnBehaVR follows FAIR principles (QR6) as it is properly available and documented. Indeed, we have generated an HTML documentation page for OWL OnBehaVR[11].

6. Final Considerations

Virtual reality is a technology that has been adopted for virtual exhibitions, safety training, medical simulations, and many other applications. Understanding user behavior in virtual reality is essential for improving user experience. However, user behavior in VR is complex due to its multimodal

[9] This paper focuses on conceptual validation. However, we clarify that simplified versions of OnBehaVR have already been successfully applied in empirical Human-Computer Interaction studies involving multimodal behavior analysis in a VR exhibition, supporting the practical relevance and applicability of OnBehaVR. [30, 31, 32, 33, 34]

[10] More specifically, we have tested OnBehaVR satisfiability through Hermit and Pellet reasoners on Protégé. See: https://protege.stanford.edu/.

[11] https://italojsoliveira.github.io/virtual-reality-ontology/.

nature, including gestures, gaze, and speech, which provide heterogeneous data streams from various sources. Structuring, interpreting, integrating, and analysing these data streams are major research problems. To address them, we propose OnBehaVR— a UFO-based core domain ontology for structuring multimodal user behavior in virtual reality. It provides conceptual clarification about this domain, enabling data integration and facilitating user behavior analysis. It satisfies relevant functional and quality requirements, including FAIR principles. In future work, we intend to implement OnBehaVR for data integration and analysis involving virtual exhibition data from our research experiments. In this process, we will investigate how to align UFO (gUFO [5]), the Ontology of units of Measure (OM), and the Semantic Sensor Network Ontology (SSN).

Declaration on Generative AI

The authors have not employed any Generative AI tools.

References

[1] M. Shehade, T. Stylianou-Lambert, Virtual reality in museums: Exploring the experiences of museum professionals, Applied sciences 10 (2020) 4031.

[2] D. Scorgie, Z. Feng, D. Paes, F. Parisi, T. Yiu, R. Lovreglio, Virtual reality for safety training: A systematic literature review and meta-analysis, Safety science 171 (2024) 106372.

[3] G. S. Ruthenbeck, K. J. Reynolds, Virtual reality for medical training: the state-of-the-art, Journal of Simulation 9 (2015) 16–26.

[4] G. Guizzardi, Ontology, ontologies and the "I" of FAIR, Data Intelligence 2 (2020) 181–191.

[5] J. P. A. Almeida, G. Guizzardi, T. P. Sales, R. A. Falbo, gUFO: A lightweight implementation of the unified foundational ontology (UFO), 2019. URL: http://purl.org/nemo/doc/gufo.

[6] D. Liu, Knowledge Graph Driven Conversational Virtual Museum Guide, Master's thesis, University of Twente, 2021.

[7] S. D. Costa, M. P. Barcellos, R. d. A. Falbo, Ontologies in human–computer interaction: A systematic literature review, Applied Ontology 16 (2021) 421–452.

[8] J. Sokołowski, K. Walczak, Semantic modelling of user interactions in virtual reality environments, in: Technological Innovation for Resilient Systems: 9th IFIP WG 5.5/SOCOLNET Advanced Doctoral Conference on Computing, Electrical and Industrial Systems, DoCEIS 2018, Costa de Caparica, Portugal, May 2-4, 2018, Proceedings 9, Springer, 2018, pp. 18–27.

[9] M. d'Aquin, A. Gangemi, M. d'Aquin, A. Gangemi, Is there beauty in ontologies?, Applied Ontology 6 (2011) 165–175.

[10] A. Jacobsen, et al., FAIR principles: interpretations and implementation considerations, Data Intelligence 2 (2020) 10–29.

[11] S. D. Costa, M. P. Barcellos, R. de Almeida Falbo, T. Conte, K. M. de Oliveira, A core ontology on the human–computer interaction phenomenon, Data & Knowledge Engineering 138 (2022) 101977.

[12] M. Messaoud, F. Cherif, C. Sanza, V. Gaildrat, An ontology for semantic modelling of virtual world, International Journal of Artificial Intelligence & Applications 6 (2015) 65.

[13] E. E. L. Pacheco, M. L. D. Sosa, M. J. T. Ruiz, M. E. C. Diaz-Velarde, Creating an ontology to represent qualitatively a scene in a virtual reality environment., Res. Comput. Sci. 147 (2018) 45–52.

[14] S.-J. Cha, G. Kim, C.-H. Jang, G.-S. Jo, An effective learning method in art using virtual reality and ontology, in: Proceedings of the International Conference on IT Convergence and Security 2011, Springer, 2012, pp. 425–434.

[15] M. Dragoni, C. Ghidini, P. Busetta, M. Fruet, M. Pedrotti, Using ontologies for modeling virtual reality scenarios, in: European Semantic Web Conference, Springer, 2015, pp. 575–590.

[16] C. Chokwitthaya, Y. Zhu, W. Lu, Ontology for experimentation of human-building interactions using virtual reality, Advanced Engineering Informatics 55 (2023) 101903.

[17] A. R. Hevner, S. T. March, J. Park, S. Ram, Design science in information systems research, MIS quarterly (2004) 75–105.

[18] G. Guizzardi, N. Guarino, Explanation, semantics, and ontology, Data Knowledge Engineering 153 (2024). URL: https://www.sciencedirect.com/science/article/pii/S0169023X24000491. doi:https://doi.org/10.1016/j.datak.2024.102325.

[19] D. Vrandečić, Ontology evaluation, in: Handbook on ontologies, Springer, 2009, pp. 293–313.

[20] L. O. Bonino da Silva Santos, B. dos Santos Vieira, C. H. Bernabé, FAIR FOR FOIS, https://w3id.org/FAIR-academic/fair4fois, 2024. [Online; accessed on May of 2024].

[21] D. J. Chalmers, The virtual and the real, Disputatio 9 (2017) 309–352. URL: https://doi.org/10.1515/disp-2017-0009. doi:doi:10.1515/disp-2017-0009.

[22] P. Brey, The physical and social reality of virtual worlds, Oxford University Press, United States, 2014, pp. 42–54. doi:10.1093/oxfordhb/9780199826162.013.029.

[23] G. Guizzardi, A. Botti Benevides, C. M. Fonseca, D. Porello, J. P. A. Almeida, T. Prince Sales, Ufo: Unified foundational ontology, Applied ontology 17 (2022) 167–210.

[24] M. Verdonck, F. Gailly, Insights on the use and application of ontology and conceptual modeling languages in ontology-driven conceptual modeling, in: Intl. Conf. on Conceptual Modeling, Springer, 2016, pp. 83–97.

[25] C. M. Fonseca, T. P. Sales, V. Viola, L. B. Da Fonseca, G. Guizzardi, J. P. A. Almeida, Ontology-driven conceptual modeling as a service, in: CEUR workshop proceedings, volume 2969, Rheinisch Westfälische Technische Hochschule, 2021.

[26] G. Guizzardi, et al., Grounding software domain ontologies in the Unified Foundational Ontology (UFO): The case of the ODE software process ontology., in: Ibero-American Conference on Software Engineering, 2008, pp. 127–140.

[27] W. Graterol, J. Diaz-Amado, Y. Cardinale, I. Dongo, E. Lopes-Silva, C. Santos-Libarino, Emotion detection for social robots based on nlp transformers and an emotion ontology, Sensors 21 (2021) 1322.

[28] R.-M. Rahal, S. Fiedler, Understanding cognitive and affective mechanisms in social psychology through eye-tracking, Journal of Experimental Social Psychology 85 (2019) 103842.

[29] H. Rijgersberg, M. Van Assem, J. Top, Ontology of units of measure and related concepts, Semantic Web 4 (2013) 3–13.

[30] D. J. Rikhtehgar, B. Usta, S. Wang, Ontology-based modeling for object segmentation and eye gaze data in vr art exhibitions, in: Joint Proceedings of Posters, Demos, Workshops, and Tutorials of EKAW 2024, volume 3967 of *CEUR Workshop Proceedings*, CEUR-WS, Amsterdam, The Netherlands, 2024.

[31] D. Javdani Rikhtehgar, S. Wang, S. Schlobach, D. Heylen, Ontology-based modeling of user-environment interaction in a virtual reality exhibition, in: Research Conference on Metadata and Semantics Research, Springer, 2023, pp. 241–252.

[32] D. Javdani Rikhtehgar, I. Tiddi, S. Wang, S. Schlobach, D. Heylen, Assessing the hi-ness of virtual heritage applications with knowledge engineering, in: HHAI 2024: Hybrid Human AI Systems for the Social Good, IOS Press, 2024, pp. 173–187.

[33] D. J. Rikhtehgar, S. Wang, S. Schlobach, D. Heylen, Real-time gaze awareness in conversational agents: Enhancing collaboration and personalization in vr art experiences, in: Proceedings of the 4th International Conference Series on Hybrid Human-Artificial Intelligence (HHAI2025), Pisa, Italy, 2025.

[34] D. J. Rikhtehgar, S. Wang, S. Schlobach, D. Heylen, Seeing and speaking with culture: How visitor profiles shape multimodal interaction in a vr art exhibition, in: Proceedings of the 25th ACM International Conference on Intelligent Virtual Agents (IVA 2025), ACM, 2025. Accepted paper.

A. Online Resources

The sources for the ceur-art style are available via

- GitHub,
- Overleaf template.