

A Reference Ontology for Information Disorder

Athus Cavalini
athus.cavalini@ifes.edu.br
Universidade Federal do Espírito
Santo
Vitória, Brazil

Giovanni Comarella
gc@inf.ufes.br
Universidade Federal do Espírito
Santo
Vitória, Brazil

João Paulo A. Almeida
jpalmeida@ieee.org
Universidade Federal do Espírito
Santo
Vitória, Brazil

ABSTRACT

The contemporary information ecosystem is being increasingly affected by information disorder, an overarching phenomenon encompassing misinformation, disinformation, and malinformation. In response to the growing need to systematically analyze and compare information disorder phenomena across platforms and contexts, the Information Disorder framework was proposed to conceptualize misinformation, disinformation, and malinformation, as well as the agents, messages, and phases involved in their production and circulation. Building on this framework, this paper proposes IDONT, a Reference Ontology for Information Disorder grounded in OntoUML and the Unified Foundational Ontology (UFO). IDONT models information disorder as a socio-temporal phenomenon and explicitly represents phase-dependent roles and motivations, enabling structured modeling of how agents participate in the creation, production, and distribution of this type of content. The ontology was operationalized in OWL and evaluated using automated reasoning and SPARQL queries over an instantiated case study based on the historical *War of the Worlds* radio broadcast. The results indicate that IDONT provides a semantically grounded conceptual foundation for the annotation, integration, and analytical investigation of information disorder phenomena.

KEYWORDS

Information Disorder, Reference Ontology, Disinformation, Misinformation, UFO, OntoUML

1 INTRODUCTION

The contemporary information environment is marked by large-scale information disorder, an overarching phenomenon encompassing misinformation, disinformation, and malinformation and involving different phases, agents, messages, and interpreters. To address this phenomenon, [21] propose the **Information Disorder** framework to conceptualize these dynamics while deliberately avoiding the term *fake news*, which is both conceptually inadequate and increasingly weaponized in public discourse.

In this framework, information disorder is distinguished into three types based on the dimensions of *falseness* and *harm*: (i) *misinformation*, when false information is shared

without intent to cause harm; (ii) *disinformation*, when false information is knowingly shared to cause harm; and (iii) *malinformation*, when genuine information is shared to cause harm, often by moving information intended to remain private into the public sphere [21].

Beyond categorizing content, the framework conceptualizes information disorder as a dynamic process unfolding over time, distinguishing between *phases* and *elements*. The phases describe how information disorder evolves through *creation*, *production*, and *distribution*, frequently involving different actors at each stage. The core elements include (a) *agents* and their respective motivations, (b) *messages* and their communicative characteristics, and (c) the interpretative roles assumed by participants as content is received, reframed, and propagated across social contexts.

Despite the conceptual richness of this framework, its elements and relations remain mostly described in natural language, which limits semantic interoperability and systematic analysis across heterogeneous datasets. These limitations can be mitigated by introducing a formal ontological representation as well as a corresponding knowledge-graph implementation. This has a number of implications: (i) settling a well-founded conceptualization for the information disorder domain, (ii) enabling the consistent annotation of information disorder events, agents, and messages, and (iii) supporting practical knowledge integration and automated reasoning.

With these objectives in mind, in this paper, we propose **IDONT** (Information Disorder Ontology), a well-founded reference ontology of information disorder. Rather than focusing solely on static content classification, IDONT accounts for information disorder as a structured socio-temporal phenomenon composed of episodes, phases, participating agents, and contextual motivations. The proposed ontology provides an interoperable conceptual backbone for annotating and integrating data, enabling comparative analysis of campaigns, supporting knowledge graph construction, and facilitating semantically grounded analyses of how harmful information is created, transformed, and disseminated.

The remainder of this paper is structured as follows: Section 2 discusses related work and identifies gaps in existing semantic models for the information disorder domain; Section 3 outlines the requirements elicitation and delimits the ontology scope; Section 4 presents the core architecture and modeling decisions of the IDONT reference ontology. Section 5 evaluates the model through operational SPARQL queries and a real-world case study. Finally, Section 6 delivers our final remarks and points out lines for future work.

In: Proceedings of the Brazilian Symposium on Multimedia and the Web (WebMedia'2026). Lavras/MG, Brazil. Porto Alegre: Brazilian Computer Society, 2026.

© 2026 SBC – Brazilian Computing Society.
ISSN 2966-2753

2 RELATED WORK

The formal representation of false or misleading information has been explored by several researchers, primarily focusing on detection and classification. Early efforts, such as the Misinformation Ontology (MO) proposed by [22], model misinformation through properties like type, motivation, communication channels, and evidential support to populate digital libraries and misinformation repositories. Later works extended these foundations to specific domains; for instance, [2] represented vaccine-related misinformation claims using semantic nanopublications encoded as RDF triples.

The work by [13] explored ontology-driven approaches for detecting false information by extracting subject–predicate–object triples from web content and comparing them against trusted knowledge bases using similarity metrics. Similarly, [12] investigated ontology-based strategies for rapid information verification during crises, assigning weights to ontology nodes to distinguish reliable information from rumors during emergency events. Although these approaches provide important foundations, they largely conceptualize misinformation as *static informational artifacts*, focusing primarily on the truth-value validation of content.

Recognizing the limitations of purely content-centric approaches, more recent research has incorporated contextual and social signals into semantic models. [4] proposed the *Fan-det* ontology to represent contextual features of social media news. This work was later extended by [23], who introduced the *iFaIDet* framework, integrating First-Order Logic (FOL) and Semantic Web Rule Language (SWRL) to improve reasoning in domains such as health and politics. Additionally, [6] proposed a multi-view ontology capturing multiple dimensions of fake news, including textual content, manipulated media, and audience propagation metrics.

Despite these advances, the current state of the art presents significant conceptual limitations. Most existing models correspond to *application ontologies* implemented directly in OWL or RDF, designed primarily to support text mining and machine learning pipelines. As a result, they adopt a *data-oriented perspective*, focusing on the representation of content features and platform metadata rather than on the intrinsic conceptual structure of the phenomenon itself. This focus frequently leads to category mixing, such as conflating intrinsic properties of agents with their contingent roles, or representing motivations as static attributes of messages rather than contextual states associated with agents.

Moreover, existing models generally adopt a static view of misinformation, failing to represent it as a dynamic socio-temporal process involving evolving roles, intentions, and interactions among agents. Finally, much of the literature still relies on the reductive and loosely defined notion of “fake news”, which overlooks the critical distinctions between misinformation, disinformation, and malinformation highlighted in the information disorder framework [21].

To address these limitations, we propose **IDONT**, a *Reference Ontology for Information Disorder* grounded in the Unified Foundational Ontology (UFO) and modeled using

OntoUML. Unlike prior application-oriented models, our approach follows a strict *conceptual modeling perspective*, prioritizing ontological clarity to represent information disorder as a complex socio-temporal phenomenon. **IDONT** models the phenomenon through *Episodes* composed of temporal *Phases* (creation, production, and distribution), capturing the complete lifecycle of information disorder [21]. Agent participation is explicitly represented through *events* (i.e., *AgentParticipation*), allowing entities to dynamically assume different roles (such as *Creator* or *Distributor*) and exhibit distinct motivations depending on the phase of the episode. By enforcing clear ontological distinctions and moving beyond the simplistic “fake news” paradigm, **IDONT** provides a semantically grounded foundation for analyzing the lifecycle and dynamics of information disorder campaigns.

3 METHODOLOGY

The development of **IDONT** follows the SABiO ontology engineering methodology [8] and is grounded in the Unified Foundational Ontology (UFO) [11]. Section 3.1 presents the ontological foundations, modeling language, and supporting tools adopted in this work, while Section 3.2 describes how the SABiO-based development process was instantiated for the development of **IDONT**.

3.1 Ontological Foundations

The proposed reference ontology, **IDONT**, is grounded in the Unified Foundational Ontology (UFO), a foundational ontology designed to provide ontological clarity for conceptual modeling [11]. UFO provides distinctions between categories such as objects, events, roles, relators, and qualities, which are particularly useful for modeling socio-temporal phenomena involving multiple actors and contextual participations.

As a modeling language, we adopt OntoUML [10], which incorporates the ontological distinctions of UFO into a UML-based conceptual modeling language. In this context, OntoUML is preferable to using “plain” UML or ER modeling because it: (i) enforces ontological clarity through well-founded categories; (ii) distinguishes *roles* from *types* (e.g., an *Agent* playing the contingent roles of *Creator*, *Producer*, or *Distributor* across different phases); (iii) supports the explicit modeling of context-dependent participations and socio-temporal events; and (iv) represents *events* and *objects* with explicit semantics (e.g., phases modeled as distinct events rather than mere attributes). These distinctions reduce modeling ambiguity and improve the overall quality of the conceptual model, particularly when representing complex socio-temporal phenomena such as information disorder episodes.

As supporting tooling, we use Tonto [7], a textual syntax for ontologies that offers specialized support for UFO-based constructs. This language enables the early identification of modeling problems that could otherwise go unnoticed during visual inspection. Moreover, Tonto specifications can be automatically transformed into multiple target languages, including OntoUML and OWL.

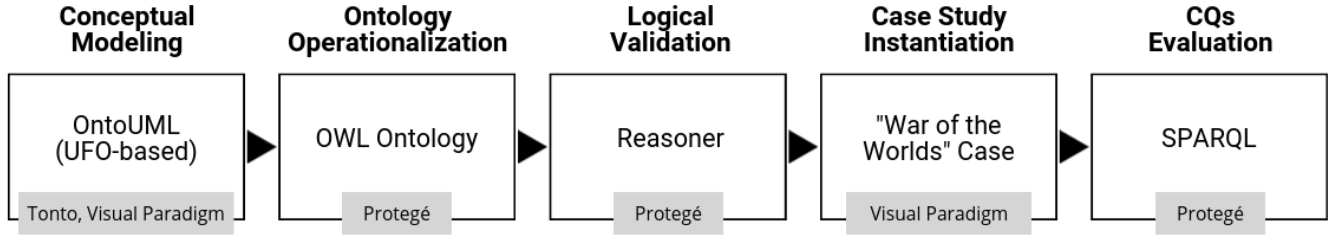


Figure 1: Overview of the ontology engineering workflow adopted in this work.

In our workflow, Tonto was initially used to experiment with and validate modeling decisions in a textual format, after which the final OntoUML model was manually refined and consolidated using Visual Paradigm [20]. The operational OWL serialization was then automatically generated from the consolidated OntoUML model. Figure 1 summarizes the ontology engineering workflow adopted in this work, spanning from initial conceptual modeling to empirical evaluation.

3.2 SABiO-based Development Process

Following SABiO [8] guidelines, the reference ontology was developed through three main activities of the ontology engineering lifecycle: domain elicitation, conceptual modeling, and formalization.

The first activity, domain elicitation, consisted of identifying the key entities, relations, and conceptual distinctions inherent to the Information Disorder framework. This step delimited core concepts such as agents, messages, disorder types, episodes, and phases, as well as the analytical dimensions used to classify information disorder phenomena.

During the conceptual modeling phase, the identified concepts were specified in OntoUML, systematically adopting the foundational ontological distinctions provided by UFO. Particular attention was given to separating rigid types (e.g., *Person*, *Organization*, *Bot*, and *Message*) from context-dependent roles and participations. Furthermore, we prioritized the explicit representation of the socio-temporal aspects of the domain through OntoUML events and context-dependent participations, treating information disorder as an emergent phenomenon resulting from agent participation rather than as an intrinsic property of messages.

Finally, the resulting conceptual model was formalized as an operational OWL ontology in order to support automated semantic querying and seamless integration with knowledge graphs. This formalization step preserves the structural and conceptual distinctions introduced in the OntoUML model while delivering machine-processable representations of information disorder episodes.

3.3 Ontology Scope

The scope of **IDONT** is the conceptual representation of the Information Disorder framework [21] as a structured socio-temporal phenomenon. Accordingly, the ontology covers: (i) the three core types of information disorder (*misinformation*,

disinformation, and *malinformation*), characterized along the analytical dimensions of *falsehood* and *harm*; (ii) the phases through which messages and content evolves (*creation*, *production*, and *distribution*), modeled as temporally bounded events composing information disorder episodes; and (iii) the main entities and relations involved in the phenomenon, including *Agent*, *Message*, and the context-dependent participations through which agents assume different roles and motivations across phases.

Rather than modeling interpretative actors as independent entity types, interpretative behavior is represented through specific participation roles assumed by agents within designated phases of an episode. This architectural decision emphasizes that participation roles are context-dependent, while agents' motivations may manifest differently across participation events.

Out of scope are algorithmic detection and classification mechanisms, platform-specific ranking or recommendation models, and low-level data structure representations. Instead, the focus of **IDONT** is on providing a domain-independent conceptual foundation that supports the semantic annotation, knowledge integration, and lifecycle analysis of information disorder episodes and their core dynamics.

These requirements were formalized as Competency Questions (CQs) that reflect the analytical lookups a semantic web system should execute over an information disorder database:

- CQ1:** How can an information disorder event be classified as *misinformation*, *disinformation*, or *malinformation*?
- CQ2:** Which agents participate in each phase of an information disorder episode?
- CQ3:** Which content formats (e.g., text, audio, video) are associated with the message artifacts involved in each phase of the episode?
- CQ4:** Which message artifact is reused across the phases of the episode?
- CQ5:** What motivations characterize agents' participation in information disorder events?

4 REFERENCE ONTOLOGY

This section reports the resulting reference ontology, **IDONT**, and the main modeling decisions derived from the scope defined in Section 3. Figure 2 summarizes the current version of the model with an OntoUML diagram. The operational

version of **IDONT** is published with the official prefix `idont:` under a permanent IRI identifier.¹ Its complete documentation, conceptual mappings, OWL serialization, and the execution artifacts used during the evaluation stage are publicly available online at the project repository.²

4.1 Core Agents and Roles

The class *Agent* is modeled as a *category*, representing entities capable of actively participating in information disorder episodes. Agents may possess a *name* attribute and are specialized into three mutually disjoint kinds: *Person*, *Organization*, and *Bot*. These kinds capture distinct essential ontological types of participating entities.

Additionally, agents may be contingently classified by OntoUML *role mixins*, namely *OfficialAgent* and *UnofficialAgent*. These role mixins capture contingent classifications that may apply to agents of different kinds depending on whether participation occurs under formal institutional authority or in informal contexts.

4.2 Communication and Messages

The class *Message* is modeled as a *kind* representing communicative content circulating within heterogeneous information environments. Its basic descriptive properties include *content*, its representational *format*, and whether it is *basedOnReality*.

Unlike previous conceptualizations that classify a message itself as misinformation, disinformation, or malinformation, IDONT treats these categories as events emerging from an agent's participation with a message. Consequently, the same message may participate in different information disorder events depending on the participants' intentions, motivations, and contextual circumstances.

This design reflects the assumption that information disorder is not an intrinsic property of a message but rather a socio-temporal phenomenon resulting from how agents create, interpret, amplify, or distribute that message.

4.3 Information Disorder Types and Classification

The abstract event *InformationDisorderEvent* represents occurrences of information disorder and is specialized into the mutually exclusive event types *MisinformationEvent*, *DisinformationEvent*, and *MalinformationEvent*. Each event is composed of one or more *AgentParticipation* instances and occurs within an *InformationDisorderEpisode*, allowing the ontology to capture the temporal and social dynamics of information disorder phenomena.

In IDONT, the attributes *basedOnReality* (*BR*), associated with *Message*, and *intentionToHarm* (*IH*), associated with *AgentParticipation*, constitute the analytical dimensions used to characterize information disorder events. The event type is determined by the logical combination of these two dimensions, according to the following rules:

$$IH \wedge \neg BR \rightarrow \text{DisinformationEvent}, \quad (1)$$

$$\neg IH \wedge \neg BR \rightarrow \text{MisinformationEvent}, \quad (2)$$

$$IH \wedge BR \rightarrow \text{MalinformationEvent}. \quad (3)$$

The combination in which *basedOnReality* = *true* and *intentionToHarm* = *false* represents legitimate information exchange and therefore falls outside the scope of information disorder in the proposed ontology.

4.4 Socio-Temporal Dynamics

The temporal lifecycle of information disorder is captured by modeling *InformationDisorderEpisode* as an *event* representing a temporally delimited episode or campaign. Episodes are identified through a *name* property, and each instance is mereologically composed of one or more *EpisodePhase* events connected via a *partOf* relation.

The *EpisodePhase* is modeled as an *event* representing a temporally bounded phase within an episode. Phases include *startDate* and *endDate* attributes and are specialized into *CreationPhase*, *ProductionPhase*, and *DistributionPhase*. Modeling phases as events rather than states enables the explicit representation of strict temporal boundaries and phase-specific participations. Each phase may comprise one or more *AgentParticipation* instances, allowing the ontology to represent the evolution of information disorder throughout an episode.

4.5 Participation and Motivation

Participation is modeled explicitly through the *AgentParticipation event*, which represents an *Agent* performing a participation involving a *Message* within a specific *EpisodePhase*. This modeling choice makes participation context-dependent and enables the same agent to interact with the same message in different phases or episodes while assuming distinct participation types, motivations, and intentions.

One or more *AgentParticipation* instances constitute an *InformationDisorderEvent*, which represents a concrete occurrence of information disorder. This design separates persistent entities (agents and messages) from the socio-temporal events that emerge from their interactions, allowing the same message to participate in different information disorder events depending on the participating agents, their intentions, motivations, and contextual circumstances.

Each participation is characterized by a *participationType*, whose values are given by the *ParticipationType* enumeration (*Creator*, *Producer*, *Distributor*, *Amplifier*, *Interpreter*). Participation types are modeled as contextual properties emerging from participation instances rather than intrinsic properties of agents.

In the original framework, "interpreter" is considered a distinct element of information disorder. In IDONT, interpreters are modeled as agents playing a specific participation role, since interpreting and reacting to a message are considered a type of participation in an episode.

¹<https://purl.org/dsl/idont#>

²<https://github.com/dsl-ufes/idont/>

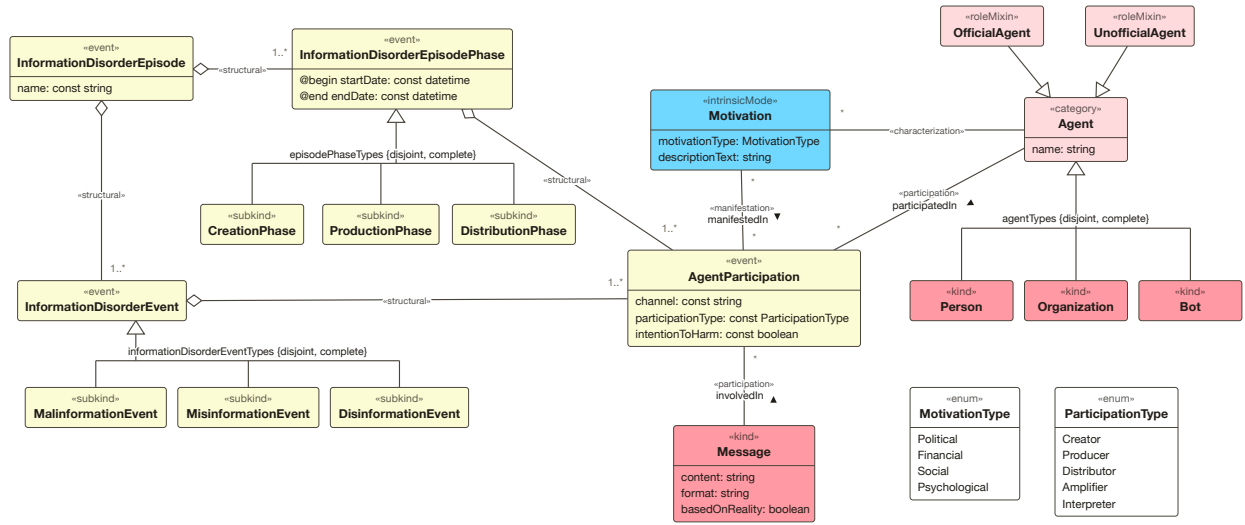


Figure 2: Reference ontology for the Information Disorder framework.

Motivational aspects are captured through the *Motivation intrinsic mode*, which characterizes an *Agent* and may be manifested in specific *AgentParticipation* events. A motivation is classified by a *motivationType*, whose values are given by the *MotivationType* enumeration (*Political*, *Financial*, *Social*, *Psychological*), and may include a descriptive specification via *descriptionText*. This design allows different motivations of the same agent to manifest across distinct participations.

Besides motivation, each participation is characterized by the agent’s *intentionToHarm*, indicating whether the agent performs that participation with the intention of causing harm. Unlike *basedOnReality*, which is associated with the message, intention is modeled as a contextual property of participation, since different agents may interact with the same message with different intentions.

4.6 Ontological Relations

The ontology includes the following main relations:

- **participatedIn** (*Agent–AgentParticipation*): each agent participation is associated with exactly one participating agent;
- **involvedIn** (*Message–AgentParticipation*): each message may be involved in zero or more participation events;
- **partOf** (*EpisodePhase–InformationDisorderEpisode*): phases compose an episode;
- **partOf** (*InformationDisorderEvent–InformationDisorderEpisode*): events occur within an episode;
- **partOf** (*AgentParticipation–EpisodePhase*): participations occur within a specific phase;
- **partOf** (*AgentParticipation–InformationDisorderEvent*): one or more participations constitute an information disorder event;

- **manifestedIn** (*Motivation–AgentParticipation*): motivations may manifest in specific participation events;

4.7 Modeling Decisions and Ontological Commitments

This section briefly discusses the main ontological commitments underlying the proposed reference ontology.

Episodes as structured events: Information Disorder phenomena are modeled as temporally delimited episodes composed of phase events. Instead of representing phases as simple enumerated states, *EpisodePhase* was modeled as an *event* specialized into *CreationPhase*, *ProductionPhase*, and *DistributionPhase*. This decision reflects the temporal and participatory nature of these phases, allowing each phase to have its own participants, temporal boundaries, and contextual properties.

Although named “Phases”, these elements were not modeled using the OntoUML «phase» stereotype. In OntoUML, phases represent intrinsic state changes of the same individual over time. In contrast, the proposed phases correspond to temporal parts of an episode, each constituting a distinct event with its own participations and relations. Therefore, they are more appropriately modeled as events connected through a mereological structure.

Contextual participation as an event: Agent involvement is modeled through the *AgentParticipation* event, which is performed by an *Agent*, involves a *Message*, and occurs within a specific *EpisodePhase*. This design makes participation context-dependent and allows the same agent to interact with the same message in different phases or episodes while assuming different participation types, motivations, and intentions.

Motivation as an intrinsic mode manifested in an agent participation: Motivational aspects are modeled as *modes* of an agent which are manifested in the context of a particular agent participation. Such modeling enables different motivations to be associated with the same agent across different phases and episodes.

Ontological differentiation of agent nature and roles: *Agent* is modeled as a *category* specialized into kinds *Person*, *Organization*, and *Bot*, capturing essential ontological distinctions between types of participating entities. In contrast, *OfficialAgent* and *UnofficialAgent* are modeled as *role mixins*, reflecting contingent relational properties of persons, organizations and bots. This separation follows core UFO distinctions, and enables orthogonal classification of agents according to different criteria.

Ontological status of information disorder types: Rather than modeling misinformation, disinformation, and malinformation as intrinsic classifications of messages, IDONT conceptualizes them as specialized information disorder events emerging from agent participations involving messages. This decision reflects the assumption that information disorder is a socio-temporal phenomenon resulting from the interaction between agents, messages, intentions, and context. Consequently, the same message may participate in multiple information disorder events throughout different phases or episodes without changing its ontological identity.

Use of enumerations: Enumerations such as *ParticipationType* and *MotivationType* were introduced as value spaces for attributes. They are not treated as ontologically independent entities, but rather as controlled vocabularies supporting conceptual clarity and operationalization.

5 EVALUATION

5.1 Ontology Instantiation

Following the development and validation of IDONT in OntoUML, the operational ontology was automatically generated in OWL using the Visual Paradigm OntoUML plugin with gUFO [1] alignment. The export was performed in Turtle serialization, mapping *idont*: classes directly to foundational *gufo*: constructs and preserving the ontological distinctions defined in the conceptual model, including categories, kinds, role mixins, events, intrinsic modes, and enumerated value spaces.

The resulting ontology was imported into Protégé [15] for validation and instantiation. No structural inconsistencies were detected after transformation, and the alignment with gUFO ensured semantic compatibility with foundational ontological commitments.

Illustrative Case

Recent semantic datasets, such as FakeNewsNet [19] and PolitiFact [14], focus primarily on message veracity and isolated text claims, lacking explicit representations of agents, motivations, and phase distinctions required by the framework. For this reason, to illustrate the analytical capabilities of IDONT, we instantiated a historical case widely regarded

as an early example of large-scale misinformation: the 1938 radio broadcast of *The War of the Worlds* [18].

The instantiation is a conceptual simplification intended for readability, while the SPARQL query results reflect the full logical structure from OWL.

The case was modeled as a single *InformationDisorderEpisode*, named *WarOfTheWorldsBroadcast*. The episode comprises three phases:

- (1) *CreationPhase*: the broadcast script was initially conceived and written as a fictional adaptation, with *OrsonWelles* participating as a *Creator*;
- (2) *ProductionPhase*: the content was adapted into a radio dramatization by *CBSRadio*, which participated as a *Producer*; and
- (3) *DistributionPhase*: the program was disseminated through radio broadcasting by *MercuryTheatreOnTheAir*, which participated as a *Distributor*.

Two *Message* instances represent the communicative artifacts involved in the episode: the broadcast script, in text format, and the broadcast itself, in audio format. In the operational knowledge base, these messages have *basedOnReality* = *false* are involved in a *MisinformationEvent*, with *intentionToHarm* = *false*, reflecting the absence of harmful intent despite misleading audience interpretation.

Multiple agents participate across phases through distinct *AgentParticipation* instances:

- *OrsonWelles*, modeled as a *Person* and classified as *UnofficialAgent*, participates in the creation phase as *Creator*, motivated by social and psychological factors related to fictional storytelling and dramatic expression;
- *CBSRadio*, modeled as an *Organization* and classified as *OfficialAgent*, participates in the production phase as *Producer*, motivated by financial and social objectives linked to broadcasting reach; and
- *MercuryTheatreOnTheAir*, also modeled as an *Organization* and *OfficialAgent*, participates in the distribution phase as *Distributor*, associated with social and psychological motivations linked to public engagement and entertainment.

Motivations are modeled as intrinsic modes of agents that manifest in specific participation instances, allowing distinct motivational contexts to coexist within the same episode.

Ultimately, this empirical instantiation demonstrates the core capacity of IDONT to systematically formalize:

- (1) the distinction between an information disorder episode and its constituent phases;
- (2) phase-dependent participation of agents assuming different contextual roles;
- (3) heterogeneous motivations associated with different stages of content circulation; and
- (4) the classification of misinformation-type events independent of malicious intent.

The reconstruction of this historical event illustrates how IDONT supports a robust conceptual foundation for the

domain, enabling semantically grounded analysis of how misleading information emerges, mutates, and propagates across complex socio-technical environments.

5.2 Competency Question Evaluation

To assess whether the operational ontology satisfies the competency questions defined during requirements elicitation, a set of SPARQL queries was executed over the instantiated knowledge base, detailed in Section 5.1. Tables 1–5 present the queries and excerpts of their results, demonstrating that IDONT operationally supports the domain analysis.

CQ1: Event classification

The ontology supports the classification of the broadcast episode according to the three types of information disorder defined in the framework. The query in Table 1 confirms that the *War of the Worlds* broadcast is correctly typed as misinformation, reflecting the absence of harmful intent while still capturing its misleading character.

Table 1: CQ1 – Event classification

Query: SELECT DISTINCT ?event ?class WHERE { VALUES ?class { :MisinformationEvent :DisinformationEvent :MalinformationEvent } ?event rdf:type ?class . }
Response: War of the Worlds broadcast -> MisinformationEvent

CQ2: Agent participation across phases

Agent participation in each phase of an information disorder episode is represented through the *AgentParticipation* event, which is performed by an *Agent* and involves a *Message* within a specific *EpisodePhase*. Since phases and participations are modeled as events, the ontology allows identifying which agents participate in each phase of an episode and which message artifacts are involved in those phases. The query presented in Table 2 confirms that participation is event-dependent, effectively validating the ontology’s representation of phase-specific involvement.

Table 2: CQ2 – Agent participation across phases

Query: SELECT DISTINCT ?phase ?agent WHERE { ?phase :partOf_1 :WarOfTheWorldsBroadcastEpisode . ?phase rdf:type ?phaseType . VALUES ?phaseType { :CreationPhase :ProductionPhase :DistributionPhase } ?participation :partOf ?phase . ?agent :participatedIn ?participation . }
Response: Creation phase -> Orson Welles Production phase -> CBS Radio Distribution phase -> Mercury Theatre on the Air

CQ3: Content formats associated with the message artifacts involved in each phase of the episode

The ontology supports retrieving the content formats associated with the message artifacts linked to each phase of the

episode. In the instantiated case, the creation phase involves the broadcast script in *text* format, while the production and distribution phases involve the broadcast message in *audio* format. This result shows that IDONT can distinguish between message artifacts and their formats across different phases of the same episode, enabling fine-grained analysis of how communicative content is transformed over time.

Table 3: CQ3: Content formats associated with the message artifacts involved in each phase of the episode

Query: SELECT DISTINCT ?phase ?message ?format WHERE { ?phase :partOf_1 :WarOfTheWorldsBroadcastEpisode . ?participation :partOf ?phase . ?message :involvedIn ?participation . ?message a :Message ; :format ?format . }
Response: Creation phase -> Broadcast script -> text Production phase -> War of the Worlds broadcast -> audio Distribution phase -> War of the Worlds broadcast -> audio

CQ4: Message artifact reused across the phases of the episode

The ontology also supports tracing whether the same message artifact is reused across multiple phases of the episode. In the instantiated case, the *War of the Worlds broadcast* appears in both the production and distribution phases, whereas the *broadcast script* is associated with the creation phase. This result demonstrates that IDONT can represent the reuse and propagation of message artifacts throughout the lifecycle of an information disorder episode.

Table 4: CQ4 – Message artifact reused across phases

Query: SELECT DISTINCT ?phase ?message WHERE { ?phase :partOf_1 :WarOfTheWorldsBroadcastEpisode . ?participation :partOf ?phase . ?message :involvedIn ?participation . }
Response: Creation phase -> Broadcast script Production phase -> War of the Worlds broadcast Distribution phase -> War of the Worlds broadcast

CQ5: Motivations associated with participation

Motivations are explicitly modeled through the *Motivation* intrinsic mode, which characterizes an agent and may be manifested in specific instances of *AgentParticipation*. Each motivation is systematically classified using the *Motivation-Type* enumeration, covering key analytical dimensions such as *Political*, *Financial*, *Social*, and *Psychological* categories. This architectural design enables semantic search applications to isolate and analyze the shifting motivations associated with specific participations within an information disorder episode. The operational consistency of this mapping is confirmed by the successful retrieval of coexisting motivational profiles, including social, financial, and psychological drivers, within the instantiated case.

These aggregated evaluation results indicate that the operational graph of IDONT successfully satisfies its complete set of defined competency questions, delivering robust and

Table 5: CQ5 – Motivations associated with participation

Query: <pre>SELECT DISTINCT ?agent ?motivation ?motivationType WHERE { ?phase :partOf_1 :WarOfTheWorldsBroadcastEpisode . ?participation :partOf ?phase . ?agent :participatedIn ?participation . ?motivation :manifestedIn ?participation ; :motivationType ?motivationType . }</pre>
Response: Orson Welles -> social, psychological CBS Radio -> financial, social Mercury Theatre on the Air -> social, psychological

executable support for structured analytical queries over empirical information disorder data. Collectively, these results demonstrate that the ontology supports campaign-oriented semantic analyses by integrating temporal, behavioral, and contextual dimensions of information disorder, enabling the reconstruction of complete episodes rather than the retrieval of isolated entities.

5.3 Potential Practical Applications

In web environments, mitigating information disorder requires moving beyond the lexical analysis of isolated messages. Recent research increasingly combines textual, behavioral, relational, and temporal evidence to understand how misleading content is created, propagated, and amplified across online platforms. By providing a structured socio-temporal semantic model, IDONT offers foundational support for the following classes of web applications.

Fact-Checking Integration: While existing fact-checking datasets provide valuable annotations regarding message veracity, they rarely capture the relationships among actors, motivations, participation roles, and dissemination processes. By leveraging the *AgentParticipation* and *Motivation* modules of IDONT, web platforms can ingest data from behavior-driven discovery pipelines—such as algorithms that identify fake news websites based on user sharing patterns [3]. This semantic enrichment enables systems to identify recurring actors, coordinated behaviors, and dissemination strategies spanning multiple information disorder episodes.

Socio-Temporal Ingestion Systems: Social media streams continuously evolve as messages are created, modified, and redistributed. Rather than treating posts as isolated observations, monitoring architectures can exploit IDONT’s event model (*InformationDisorderEpisode* and *EpisodePhase*) to represent the temporal evolution of campaigns. This semantic representation supports the identification of diffusion trajectories, amplification patterns, and malicious triggers [17], enabling richer analyses of campaign dynamics over time.

Cross-Platform Campaign Analysis: Information disorder campaigns rarely remain confined to a single platform. Instead, narratives, actors, and media assets migrate across social networks, messaging applications, fact-checking repositories, and monitoring systems, creating heterogeneous datasets with fragmented terminologies and isolated metadata [9, 16]. IDONT provides a standardized, gUFO-aligned semantic vocabulary capable of integrating these independent sources into a unified knowledge graph. By explicitly representing agents,

participation roles, motivations, messages, and episode phases, the ontology supports semantic querying and reasoning across platforms, enabling the reconstruction of coordinated campaigns, the identification of recurrent dissemination patterns, and interoperability among monitoring, detection, and mitigation systems [5].

Overall, IDONT complements existing misinformation detection approaches by providing a shared semantic foundation capable of integrating heterogeneous data sources and supporting campaign-oriented analyses that extend beyond isolated message classification.

6 FINAL REMARKS

This paper presented a reference ontology for the Information Disorder framework grounded in OntoUML and the Unified Foundational Ontology (UFO). The proposed model conceptualizes information disorder as a structured socio-temporal phenomenon composed of episodes, phases, messages, agents, and context-dependent participations. By explicitly representing phase-dependent participations, roles, and motivations, the ontology captures the dynamic and multi-actor nature of information disorder beyond content-centered classifications.

The ontology was operationalized in OWL and validated through the instantiation of the historical *War of the Worlds* radio broadcast case. The evaluation demonstrated its ability to consistently represent heterogeneous agents, motivations, and participation roles while supporting executable competency-question queries for classification, participation analysis, semantic traceability, and message propagation across episode phases. By modeling each phase as an event linked to messages and agent participations, the ontology also enables the reconstruction of information disorder episodes, allowing analysts to trace how campaigns originate, evolve, and propagate across actors and contexts.

Overall, IDONT provides a conceptual foundation that may support the semantic annotation, integration, and analysis of heterogeneous misinformation and disinformation datasets. By supporting semantic interoperability, knowledge graph construction, and campaign-oriented analyses, the ontology establishes a reusable semantic layer for future research on information disorder across platforms.

As future work, we plan to validate the ontology using contemporary social media datasets, investigate its alignment with existing misinformation and media ontologies, and explore its integration with automated analysis pipelines and large-scale knowledge graph infrastructures.

Ethical Care

This work complies with the Brazilian CNS Resolution N.º 510/2016.

Acknowledgments

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES, Finance Code 001), the Espírito Santo Research Foundation (FAPES, #2022/NGKM5, #2021/GL60J, #2025-1H3FP, #2026-B74MN, #2026-PPX8T) and CNPq (443130/2023-0, 313412/2023-5). Generative AI tools were used to assist with grammar and spelling corrections. No generative AI were used in the analysis or interpretation of results. The authors take full responsibility for the content of the publication.

REFERENCES

- [1] João Paulo A. Almeida, Giancarlo Guizzardi, Tiago Prince Sales, and Claudenir M. Fonseca. 2026. gUFO: A Gentle Foundational Ontology for Semantic Web Knowledge Graphs. arXiv:2603.20948 [cs.AI] <https://arxiv.org/abs/2603.20948>
- [2] Muhammad Amith and Cui Tao. 2018. Representing vaccine misinformation using ontologies. *Journal of Biomedical Semantics* 9, 22 (2018), 1–13.
- [3] Leandro Araujo, João Couto, Luiz Nery, Isadora Rodrigues, Jussara Almeida, Julio Reis, and Fabricio Benevenuto. 2024. Finding Fake News Websites in the Wild. In *Proceedings of the 30th Brazilian Symposium on Multimedia and the Web* (Juiz de Fora/MG). SBC, Porto Alegre, RS, Brasil, 171–178. <https://doi.org/10.5753/webmedia.2024.243115>
- [4] Anoud Bani-Hani, Oluwasegun Adedugbe, Elhadj Benkhelifa, Munir Majdalawieh, and Feras Al-Obeidat. 2020. A semantic model for context-based fake news detection on social media. In *2020 IEEE/ACS 17th International Conference on Computer Systems and Applications (AICCSA)*. IEEE, 1–7.
- [5] Athus Cavalini, Thamya Donadia, Fabio Malini, and Giovanni Comarela. 2024. Detecting Misinformation on Telegram Antivaccine Communities. In *Anais do XXXIX Simpósio Brasileiro de Bancos de Dados* (Florianópolis/SC). SBC, Porto Alegre, RS, Brasil, 729–735. <https://doi.org/10.5753/sbbd.2024.243188>
- [6] K Chandrasekaran, A Kandasamy, M Venkatesan, P Prabhavathi, M Gokuldeh, and C Aishwarya. 2023. Ontology for Contextual Fake News Assessment based on Text and Images. In *Preprint / Working Paper*.
- [7] Matheus Lenke Coutinho. 2023. *Tonto: A Textual Language for Ontology-Driven Conceptual Modeling*. Master's thesis. Universidade Federal do Espírito Santo (UFES), Centro Tecnológico, Colegiado do Curso de Ciência da Computação, Vitória, ES.
- [8] Ricardo de Almeida Falbo. 2014. SABiO: Systematic Approach for Building Ontologies. In *Proceedings of the Ontology and Conceptual Modeling Research Group Workshop (OntoComODISE 2014)* (CEUR Workshop Proceedings, Vol. 1301). 1–14. https://ceur-ws.org/Vol-1301/ontocomodise2014_2.pdf
- [9] João Guilherme Bastos dos Santos and Tatiana Maria Dourado. 2020. Dissemination of Fake News in Cross-Platform Perspective: The Hydra of 2018 Brazilian Presidential Election. In *Proceedings of the #AoIR 2nd Flashpoint Symposium 2020*. Galoá, Porto Alegre, Brazil (Online). <https://proceedings.science/aoir-2nd/trabalhos/dissemination-of-fake-news-in-cross-platform-perspective-the-hydra-of-2018-brazi?lang=en>
- [10] Giancarlo Guizzardi. 2005. *Ontological foundations for structural conceptual models*. PhD Thesis - Research UT, graduation UT. University of Twente.
- [11] Giancarlo Guizzardi, Alessander Botti Benevides, Claudenir M. Fonseca, Daniele Porello, João Paulo A. Almeida, and Tiago Prince Sales. 2022. UFO: Unified Foundational Ontology. *Applied Ontology* 17, 1 (2022), 167–210. <https://doi.org/10.3233/AO-210256>
- [12] Chelsea Hicks. 2017. An Ontological Approach to Misinformation: Quickly Finding Relevant Information. In *Hawaii International Conference on System Sciences (HICSS)*. URI: <http://hdl.handle.net/10125/41263>.
- [13] Svitlana Mazepa, Sergiy Pugach, Serhiy Banakh, Oleksandra Yavorska, Andriy Melnyk, and Natalia Golota. 2021. An Ontological Approach to Detecting Fake News in Online Media. In *2021 11th International Conference on Advanced Computer Information Technologies (ACIT)*. IEEE, 531–535.
- [14] Rishabh Misra and Jigyasa Grover. 2022. *Do Not 'Fake It Till You Make It'! Synopsis of Trending Fake News Detection Methodologies Using Deep Learning*. 213–235. https://doi.org/10.1007/978-3-031-10869-3_12
- [15] Mark A. Musen. 2015. The protégé project: a look back and a look forward. *AI Matters* 1, 4 (2015), 4–12. <https://doi.org/10.1145/2757001.2757003>
- [16] João Olivetti, Ana Bomfim, Marcos Oliveira, Helom Marques, Juan Avelar, Julio Reis, and Philipe Melo. 2025. O Fenômeno da Desinformação no WhatsApp, Telegram e Outras Plataformas de Mensagens Instantâneas: Uma Revisão Sistemática da Literatura. In *Proceedings of the 31st Brazilian Symposium on Multimedia and the Web* (Rio de Janeiro/RJ). SBC, Porto Alegre, RS, Brasil, 661–675. <https://doi.org/10.5753/webmedia.2025.16058>
- [17] Marcelo Santos, Chang Sup Park, João Guilherme Bastos dos Santos, Athus Cavalini, and Homero Gil de Zúñiga. 2025. Latent bombs of disinformation. *Information, Communication & Society* 28, 16 (2025), 2868–2889. <https://doi.org/10.1080/1369118X.2025.2492575>
- [18] Brad Schwartz. 2018. Orson Welles and the Birth of Fake News. *The New York Times* (30 October 2018). <https://www.nytimes.com/2018/10/30/opinion/orson-welles-war-of-the-worlds-fake-news.html> Accessed: 2026-02-26.
- [19] Kai Shu, Deepak Mahudeswaran, Suhang Wang, Dongwon Lee, and Huan Liu. 2018. FakeNewsNet: A Data Repository with News Content, Social Context and Dynamic Information for Studying Fake News on Social Media. *arXiv preprint arXiv:1809.01286* (2018).
- [20] Visual Paradigm. 2026. *Visual Paradigm*. <https://www.visual-paradigm.com/> Computer software.
- [21] Claire Wardle and Hossein Derakhshan. 2017. Information disorder: Toward an interdisciplinary framework for research and policy making. <https://rm.coe.int/information-disorder-report-november-2017/1680764666>
- [22] Lina Zhou and Dongsong Zhang. 2007. An Ontology-Supported Misinformation Model: Toward a Digital Misinformation Library. *IEEE Transactions on Systems, Man, and Cybernetics—Part A: Systems and Humans* 37, 5 (2007), 804–813.
- [23] Hussaini Abubakar Zubairu and Enesi Femi Aminu. 2024. Towards an Improved Semantic Model: An Ontological Based Framework for Misinformation Detection. *SSRN* (2024).