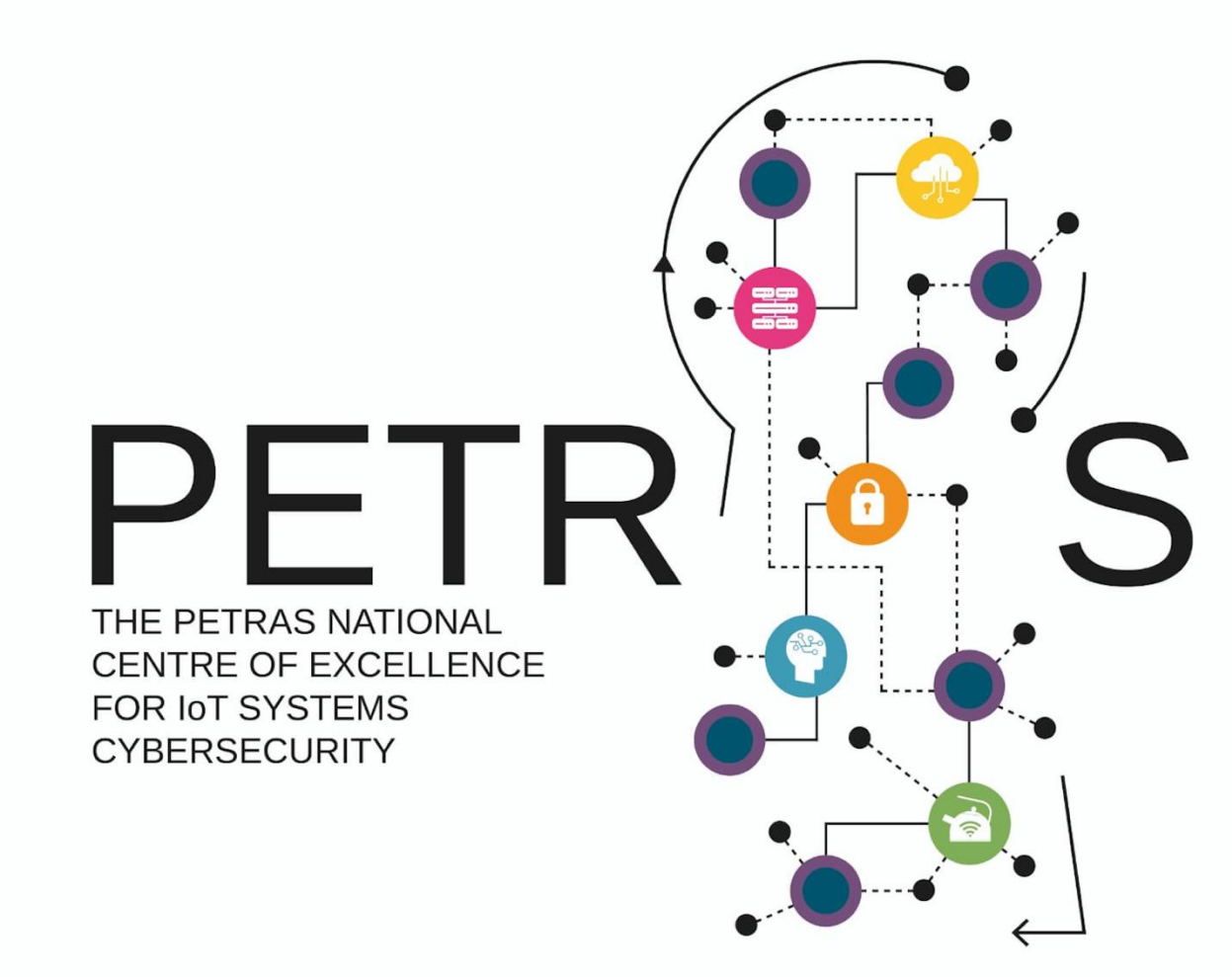

SOfIoTS: Ontological Framework, Demonstration Outcomes, and Recommendations for Further Work (D3/D4)

17 July 2023 (version 1.0)

Authors: Paul Smart, Michael Boniface, Jarwar Aslam
and Jeremy Watson CBE

Project Title: Secure Ontologies for Internet of Things Systems (SOfIoTS)

Details of document preparation and issue:

Version no.	Prepared	Checked	Reviewed	Approved	Issue date	Issue status
1.0	Paul Smart				17/07/23	Draft

CONTENTS

LIST OF TABLES

EXECUTIVE SUMMARY

The Secure Ontologies for the Internet of Things (SOfIoTS) project seeks to advance our understanding of the current state-of-the-art in respect of security provision in IoT ontologies. It also aims to extend the current state-of-the-art by specifying an expansible IoT ontological framework that can be integrated with the UK Digital Twin model. The present report summarizes the progress made in respect of these objectives. In particular, we describe how a common, upper-level ontology, called Basic Formal Ontology (BFO), can be used to model security concepts, IoT devices, digital twins, IoT data flows, and human factors. While BFO is not the only upper-level ontology that could be used for IoT security modelling, there are a number of reasons that make it a compelling choice for the SOfIoTS project. Aside from the fact that BFO is one of the most widely used upper-level ontologies, it also serves as the basis for the Industrial Ontologies Foundry (IOF). BFO also serves as the basis for a prominent cyber ontology initiative that subsumes work by The MITRE Corporation. Finally, there have been a number of attempts to apply BFO to Building Information Modeling (BIM), and BFO was one of the upper-level ontologies surveyed as the part of the effort to develop an Information Management Framework (IMF) for the UK National Digital Twin (NDT) initiative.

Prior work has identified a number of recurring concepts across security ontologies. These include the concepts of threat, risk, vulnerability, asset, security mechanism, and so on. In the present report, we discuss how each of these concepts can be situated within a BFO-conformant ontology. As far as we are aware, this is the first attempt to provide an ontological characterization of security-related concepts from a BFO perspective.

In addition to security concepts, we also discuss how IoT devices, digital twins, and IoT information flows can be represented in BFO. Again, as far as we are aware, this represents the first attempt to apply BFO to the realm of IoT devices and Cyber-Physical Systems.

Finally, we explore how BFO could be applied to the modelling of human factors, focusing specifically on the notion of capabilities. We also outline an ontological approach to the representation of trust-related concepts, drawing on research that is spread across a number of disciplines, including sociology and analytic philosophy.

The present report makes a number of substantive contributions to the field of security modelling and IoT ontologies. These contributions include the following:

- A mapping of the W3C Semantic Sensor Network (SSN) ontology to a mid-level extension of BFO.
- An ontological approach to the modelling of IoT data flows.
- A novel account of value that is inspired by recent work in cognitive neuroscience and generative AI.
- A BFO-conformant approach to the representation of trust and trustworthiness
- A BFO-conformant approach to the representation of digital twins.
- An innovative proposal regarding the use of biophysical principles to inform the design of future Cyber-Physical Systems.

The report concludes with a set of recommendations pertaining to future work. These recommendations are grouped under the following headings:

1. **Ontology Design Principles.** We recommend that future ontologies should adhere to a set of design principles to support reuse, interoperability, and human comprehension.

2. **Modularity.** We recommend that future IoT ontologies (including those for IoT security) be developed in a modular fashion, such that each ontology focuses on a restricted range of terms and concepts.

3. **Modelling Patterns.** We recommend that future work should strive to produce a library of modelling patterns that can be used to guide and inform practical efforts to model IoT systems. Such a library should rely on concrete examples to show how BFO can be applied to the modelling of (e.g.) IoT data flows and the security-hazards that emerge in respect of such flows.

4. **Ontology Repositories.** We recommend the creation of an Internet Ontology Portal to facilitate the communal development, dissemination, and application of Internet-related ontologies. This portal should emulate the functionality exhibited by portals that are already in use for the biomedical domain (e.g., https://bioportal.bioontology.org/).

5. **Trust, Trustworthiness, and Human Factors.** We recommend that future work should build on the present work by developing dedicated ontologies for trust, trustworthiness, and human factors.

6. **Philosophical Engineering.** Many of the challenges associated with ontology development can be traced to shortfalls in our understanding of key terms and concepts. This understanding is, however, largely available in contemporary analytic philosophy. For this reason, we recommend that future ontology-oriented work should feature more robust forms of interdisciplinary collaboration with the discipline of philosophy, especially with the fields of metaphysics and axiology.

7. **Active Inference.** Finally, we recommend that closer attention be paid to the so-called active inference framework, which has been proposed as a theoretically unified account of brain function. From an algorithmic perspective, we suggest that this framework could be useful when it comes to optimizing information flows, advancing the state-of-the-art in machine learning, and designing the next generation of Cyber-Physical Systems.

The implementation of these recommendations will, we suggest, support the effort to develop high-quality ontologies that can be used to represent the properties of IoT devices, the information flows in which these devices participate, and the wider nexus of forces and factors that influence the successful operation of future Cyber-Physical Systems.

1 INTRODUCTION

1.1 The SOfIoTS Project

The Internet of Things (IoT) refers to the vast network of interconnected devices, sensors, and systems that communicate and exchange data over the Internet. These devices range from simple everyday objects, such as thermostats and light bulbs, to complex industrial machinery and critical infrastructure components. The recent proliferation of these devices has created new opportunities for control and communication, but it has also given rise to multiple security concern. As a result, the security of IoT devices has become a pressing concern, demanding robust countermeasures to safeguard the privacy, integrity, and availability of sensitive data and critical systems.

The Secure Ontologies for the Internet of Things (SOfIoTS) project seeks to advance our understanding of the current state-of-the-art in respect of security provision in IoT ontologies. It also aims to extend the current state-of-the-art by specifying "an expansible IoT device and system ontological framework that can be integrated into the UK Digital Twin model." Additional goals include better support for standardized descriptions of IoT data sources and data flows, ontological descriptions that support machine learning functionality at the 'edge', and ontological frameworks that align with the objectives of the Digital Built Britain (DBB) initiative, particularly with respect to Building Information Modeling (BIM) and digital twins.

1.2 Scope of the Present Report

The present report summarizes our progress made in respect of these objectives. Our overarching objective is to develop a domain-level ontology that extends an upper-level (or top-level) ontology known as Basic Formal Ontology (BFO) (Arp, Smith, and Spear 2015). This domain-level ontology is intended to provide an ontological framework for IoT devices and security concepts. One of the virtues of BFO is that it has been applied to multiple domains, which provides an opportunity for integration and interoperability across seemingly disparate areas of ontological research. In addition, the ontological commitments of BFO provide a springboard for philosophical debates about the nature of security-related concepts. As far as we are aware, there has been no attempt to link BFO to the realm of digital twins; nevertheless, in the present report, we show how such support can be provided via a BFO-conformant ontology dubbed the Common Core Cyber Ontology (C3O) (Donohue et al. 2018). In respect of the alignment with BIM, there have been recent efforts to apply BFO to BIM (Park and Shin 2023). While we do not discuss these efforts here, there is, we suggest, no reason to think that BFO is incompatible with BIM-related efforts. Finally, in respect of the machine learning objectives, we show how BFO provides a standardized approach to the representation of IoT data. We do not discuss the specific mechanisms by which the resultant information could be leveraged by machine learning algorithms; nevertheless, we suggest that the use of a standardized representational scheme could be used to support the development of next-generation intelligent systems built around the use of biophysical principles, such as those associated with the active inference framework (Da Costa et al. 2022; Friston et al. 2015; Parr, Pezzulo, and Friston 2022). Recent innovations in this area suggest that such principles may provide a unified approach to understanding brain function, as well as providing a degree of explanatory unification in respect of generative Artificial Intelligence (AI) systems. What is more, recent work has identified approaches to developing explainable AI systems based around the active inference framework (see Albarracin et al. 2023).

1.3 Report Structure

The structure of the report is as follows. In Section 2, we review earlier work undertaken in respect of the SOfIoTS project (Jarwar, Tooth, and Watson 2022), as well as more general work on security ontologies. In Section 3, we provide an overview of BFO, as well as a suite of mid-level ontological extensions to BFO, known as the Common Core Ontologies (CCO). Section 4 discusses how common security concepts might be represented in BFO. Section 5 presents our approach to the representation of IoT devices within BFO. This includes our approach to the representation of information flows, the capabilities/functionality of IoT devices, and the role of IoT devices in controlling/regulating the dynamics of physical systems. Section 6 addresses one of the gaps identified in earlier work by Jarwar et al. (2022). This relates to the omission of human factors information, which is deemed crucial to our understanding of risks and mitigation strategies. Our analysis, here, is admittedly cursory in nature. Nevertheless, we provide the basis for more detailed extensions, showing how key human factors considerations can be represented in a BFO-conformant manner. Finally, Section 7 summarizes our recommendations with regard to future work. Our key proposal is that future efforts should build on the present effort by developing ontologies in a manner that conforms to the ontological distinctions made in an upper ontology. We also suggest that greater attention is required to the means by which ontologies are communicated and shared with the wider community.

Table 1. Summary of key contributions.

Contribution	Location
Mapping of the W3C Semantic Sensor Network (SSN) ontology to a mid-level extension of BFO.	Section 5.2
Analysis of the security mechanism concept from a neo-mechanical standpoint.	Section 4.8
Outline of an ontological approach to the modelling of IoT data flows.	Section 5
An ontological account of value that is inspired by recent work in cognitive neuroscience and generative AI.	Section 4.2
The first attempt to provide a BFO-conformant representation of digital twins.	Section 5.8
The first attempt to represent risk-related terms in a BFO-conformant manner.	Section 4.4
A BFO-conformant approach to the representation of trust and trustworthiness.	Section 6
An innovative proposal regarding the use of biophysical principles to inform the design of future cyber-physical systems.	Section 7.7

1.4 Contributions

The specific contributions of the present work to the field of security- and IoT-related research are discussed in a somewhat distributed manner throughout the report. As a means of bringing these contributions into sharper focus, Table 1 summarizes some of the contributions that have been made during the course of the present work effort.

1.5 Notation

Throughout the present report, we rely on diagrammatic notations to help clarify key points. These are mostly Unified Modeling Language (UML) object diagrams, which were created with the help of Visual Paradigm (Community Edition).[1]

The text makes references to ontology classes, properties, and individuals (instances). These are rendered as follows:

- Classes are rendered **`LIKE THIS`**.
- Properties (relations) are rendered ***`like this`***.
- Individuals (instances) are rendered `:LIKE THIS`.

Throughout the report, we adopt an Aristotelian approach to definition, drawing attention to the features that distinguish the definiendum from its superordinate category (see Arp et al. 2015, for further details). The general schema for an Aristotelian definition is:

> **S** =*def.* a G that Ds.

Where "G" (for genus) is the immediate parent term of "S" (for species), which is the term that is being defined. "D" stands for differentia, which is to say "D" tells us what it is about certain Gs in virtue of which they are Ss.

[1] See https://www.visual-paradigm.com/.

2 SECURITY ONTOLOGIES: THE STATE-OF-THE-ART

As noted by Jarwar et al. (2022), security ontologies "are developed to represent and standardize cybersecurity knowledge through a common vocabulary and machine-interpretable formalism." It is widely believed that such ontologies are poised to deliver a number of benefits, the more notable of which are as follows:

- **Standardization:** Security ontologies provide a common language and standardized representation for security-related concepts, terms, and relationships. This promotes interoperability and facilitates effective communication and collaboration among different security systems, tools, and stakeholders. It also helps avoid ambiguities and inconsistencies that can arise due to the diverse terminology used in the security domain.

- **Knowledge Sharing and Reuse:** Ontologies enable the sharing and reuse of security knowledge. They provide a structured and formalized way to capture and represent security domain expertise, best practices, and lessons learned. By leveraging existing ontologies, security professionals can access and integrate valuable knowledge, reducing redundancy and promoting efficiency in security-related activities.

- **Improved Threat Intelligence:** Security ontologies enhance the analysis and management of threat intelligence. They enable the integration of diverse security data sources, such as vulnerability databases, intrusion detection systems, and threat feeds, by providing a common framework to represent and reason about these different types of information. This integration facilitates better threat detection, correlation, and response capabilities.

- **Enhanced Situational Awareness:** Ontologies help in creating a comprehensive understanding of the security landscape by representing the relationships between various security entities and events. This enables the development of advanced situational awareness systems that can detect and analyse security incidents in real-time, providing security operators with a holistic view of the environment and aiding in effective decision-making.

- **Automated Reasoning and Analysis:** With the use of security ontologies, automated reasoning techniques can be applied to perform complex security analysis tasks. By encoding security rules, policies, and constraints in the ontology, automated systems can reason over security-related data and make inferences. This can aid in threat detection, risk assessment, access control, and policy enforcement, thereby augmenting human capabilities and reducing the manual effort required.

- **Scalability and Adaptability:** Security ontologies provide a scalable and adaptable framework for capturing and representing security knowledge. As the security domain evolves and new threats emerge, ontologies can be updated and extended to incorporate the latest information. This flexibility allows security systems to adapt and respond to changing circumstances, ensuring their continued relevance and effectiveness.

Given these benefits, it will come as no surprise to learn that security ontologies are a prominent focus of research attention. Indeed, there have been many efforts to develop security ontologies over the past two decades. Some of these ontologies focus on more general issues, such as the modelling of security-related concepts. Others focus on specific aspects of the security domain, such as risk

assessment, threat incident modelling, and vulnerability analysis. Unfortunately, despite the considerable interest that has been expressed in security ontologies, it remains unclear whether contemporary ontologies are suitably poised to deliver on the aforementioned benefits. According to a recent survey, there are a number of problems with contemporary security ontologies (Oliveira et al. 2021). These include the apparent unavailability of many security ontologies, the failure to align security ontologies with an upper or foundational ontology, and (relatedly) the failure to converge on a common approach to the modelling of cyber-security concepts (Oliveira et al. 2021). This latter shortcoming is, of course, somewhat ironic given that one of primary purposes of an ontology is to provide a standardized approach to knowledge representation.

Our own survey of security ontologies has revealed a number of additional shortcomings. As noted by Jarwar et al. (2022), security ontologies often overlook features that are particularly relevant to security in an IoT context. Such shortcomings include a failure to appreciate the computational and energetic constraints associated with IoT devices and a failure to adequately address socio-technical and human factors considerations.[2]

Beyond this, there are reasons to doubt the overall quality of existing security ontologies. In some cases, ontologies are little more than terminological taxonomies that make little use of the semantic axioms available in ontology languages. In other cases, ontologies deliver inaccurate or inconsistent results when it comes to the semantic characterization of certain terms. Donohue et al. (2018), for example, suggest that the Unified Cyber Ontology (UCO) is prone to produce inconsistent results given its mapping to a wider set of Semantic Web resources.

In general, contemporary security ontologies suffer from the same shortcomings as those identified for the more general realm of engineering ontologies. A nice summary of these shortcomings is provided by Hagedorn et al. (2019):

> Despite the breadth of ontologies proposed for various engineering subdomains […] their use in industry remains relatively rare. While several factors are likely culprits in this lack of uptake, many issues stem from failures of interoperability. There are just too many engineering ontologies, almost all of which are developed in an ad hoc fashion with little attention to issues of orthogonality, cross-ontology compatibility and sustainability. Few engineering ontologies utilise a top-level ontology to organise their terms, provide development guidelines, or establish a basic philosophical-architectural perspective. As a result, few engineering ontologies adhere to shared modelling principles, and so interoperability between any two engineering ontologies is rare. […] Many ontologies published in the engineering literature have not been made publicly available, meaning they provide little in the way of input to subsequent ontology development or of lessons of consequence for the construction and application of ontology-based engineering tools by subsequent generations. […] Existing engineering ontologies are often overlapping, non-interoperable, unreadable by humans, and defined in an esoteric fashion that limits their usefulness to the broader community. For simple, self-contained applications these issues may not be significant. However, few engineering applications are simple, and many

[2] This is not to say the literature has been utterly silent on these issues. In respect of human factors, for example, Oltramari et al. (2014) propose an OWL-based ontological framework that is constituted by a domain ontology of cyber operations and extended with a security-related middle-level ontology. This ontology has since been extended with a human factors ontology to support the representation of individual characteristics and the forces and factors that influence trust-related ascriptions (Oltramari et al. 2015).

> of the core advantages offered by ontologies depend precisely on formality, interoperability, availability, and usefulness beyond any single application. (Hagedorn et al. 2019, p. 629)

Part of the aim of the SOfIoTS project is to address these shortcomings via a combination of suggested solutions and recommendations pertaining to future work. As noted in Section 1, our approach relies on the use of a top-level ontology called BFO, as well as a modular suite of mid-level extensions to BFO. This, we suggest, helps to address some of the problems identified by Hagedorn et al. (2019), Oliveira et al. (2021), and Jarwar et al. (2022). Table 2 summarizes our approach to the resolution of prominent issues and concerns.

Table 2. Approach to resolving issues with contemporary IoT security ontologies.

Issue	Solution
Poor re-use of existing ontologies.	We adopt a modular approach to ontology development, drawing on a suite of existing ontologies that are used across multiple domains. In Section 7.4, we recommend that greater attention be devoted to developing a repository for sharing Internet- and Web-related ontologies, including those devoted to representing security-related concepts.
Interoperability with other ontologies.	BFO provides a common approach to the modelling of domain-relevant terms, which facilitates the mapping to other ontologies.
Lack of support for human factors considerations.	Section 6 shows how human factors considerations can be incorporated into a BFO-conformant ontology.
Limited applicability to IoT devices.	Section 5 presents a BFO-conformant approach to the representation of IoT devices. We also show how the W3C SSN ontology can be mapped to BFO.
Information modelling and data interoperability	Section 5 discusses how a BFO approach to information modelling (see CUBRC, 2020b) can be used to provide a common (standardized) approach to representing the information entities generated by, processed by, and communicated by IoT devices.
Limited support for digital twins.	Section 5.8 discusses how digital twins can be represented within BFO.
Disparate approaches to the modelling of security concepts	Section 4 seeks to advance our understanding of security concepts by discussing their taxonomic position within the BFO class hierarchy.
Poor quality of existing ontologies.	In part, this problem relates to the absence of guidelines and concrete examples demonstrating the application of BFO to common modelling problems. In Section 7.3, we suggest that one of the targets for future work is to provide a comprehensive suite of examples demonstrating the application of BFO to concrete IoT-related scenarios. The aim here is to develop a set of best-practice guidelines pertaining to the application of BFO in Web- and Internet-related contexts.

3 BASIC FORMAL ONTOLOGY

3.1 Introducing Basic Formal Ontology

As noted by Oliveira et al. (2021), one of the issues raised by a survey of extant security ontologies relates to an under-utilization of upper-level ontologies. In particular, Oliveira et al. (2021) found that of the 57 security ontologies selected for study only four made use of an upper-level ontology.

The term "upper-level" (or "top-level" or "foundational") ontology refers to an ontology that is intended to represent the most general aspects of reality, such as the distinction between continuants (or endurants) and occurrents (or perdurants). Unlike other types of ontologies, an upper-level ontology restricts its focus to entities that are common to *all* domains of discourse. In this respect, it differs from so-called mid-level or domain-level ontologies, which tend to focus on entities that are limited to particular domains. Here is how Donohue et al. (2018) characterize the distinction between upper-level, mid-level, and domain-level ontologies:

- An upper-level ontology is an ontology that represents only highly generic categories of entity (e.g., object, quality, function, process) and their relationships to each other (e.g., componential or taxonomic relationships).

- A mid-level ontology is an ontology that represents relatively general categories common to many domains of interest (e.g., person, act of communication, country).

- A domain-level ontology is an ontology that represents categories that are of interest to a more limited number of domains (e.g., intelligence analyst role, portion of ammonium nitrate, or watercraft registration).

BFO is one of a number of upper-level ontologies that have been developed to assist with ontology engineering efforts (Arp, Smith, and Spear 2015; Otte, Beverley, and Ruttenberg 2022). Some other notable upper-level ontologies, include Descriptive Ontology for Linguistic and Cognitive Engineering (DOLCE) (Borgo et al. 2022), Unified Foundational Ontology (UFO) (Guizzardi et al. 2022), General Formal Ontology (GFO) (Loebe, Burek, and Herre 2022), and Yet-another more advanced top-level ontology (YAMATO) (Mizoguchi and Borgo 2022).[3] While some of the entities described by BFO resemble those included in other upper-level ontologies, there are a number of features that distinguish BFO from other upper-level ontologies. One of these relates to the so-called core commitments of BFO. In particular, BFO is committed to the following ontological principles (Arp, Smith, and Spear 2015; Otte, Beverley, and Ruttenberg 2022):

- **Realism:** BFO is committed to ontological realism (Smith and Ceusters 2010), which mandates a concern with entities that exist in reality. Consequently, BFO consists fundamentally of representations of reality rather than merely language, concepts, or mental representations about reality.

- **Fallibilism:** BFO accepts that future research may reveal the need for an expansion or restructuring of the categories that BFO recognizes. For this reason, BFO is committed to tracking scientific developments over time, and updating ontologies in accordance with scientific developments.

[3] See Borgo et al. (2022), for an overview of upper-level ontologies.

- **Adequatism:** BFO is committed to the idea that discipline-specific entities are worthy of representation in their own right. It thus rejects the principles of reductionism, which assumes that the entities in one domain (e.g., biology) can be reduced to the entities of a more fundamental domain (e.g., physics).

One of the things that makes BFO interesting from an applied ontology perspective is that it is one of the most widely used upper-level ontologies, serving as the basis for both the Open Biological and Biomedical Ontology (OBO) Foundry (Smith et al. 2007) and the Industrial Ontologies Foundry (IOF) (Smith et al. 2019; Drobnjakovic et al. 2022). Its main application has been in the biomedical and life sciences domain, where it serves as the basis for over 350 mid-level and domain-level ontologies. These include ontologies for medical science (Scheuermann, Ceusters, and Smith 2009), infectious diseases (Babcock et al. 2021; Goldfain, Smith, and Cowell 2010), disease resistance (Goldfain, Smith, and Cowell 2011), phenotype modelling (Köhler et al. 2014; Le and Dao 2018), and plant development (Walls et al. 2019). More recently, BFO has extended its reach to areas beyond the biomedical domain. These include the domains of physics (Cheong and Butscher 2019), BIM (Park and Shin 2023), cognitive processes (Limbaugh et al. 2020), intelligence analysis (Mandrick and Smith 2022), and additive manufacturing (Hagedorn et al. 2019). This widespread usage makes BFO an interesting target for ontology development efforts in the security domain. As yet, however, there are relatively few ontologies that rely on BFO as an upper-level ontology. Oliveira et al. (2021) report that only one of the 57 ontologies they studied relied on BFO. This ontology, described by Casola et al. (2019), aims to use BFO as a means of developing a domain-level security ontology that is aimed at modelling the International Organization for Standardization (ISO) 27001 family of standards. While this effort sounds promising, it is unclear how much progress has been made on the ontology. [Unfortunately, we were unable to locate an online version of the ontology described by Casola et al. (2019).]

What we seem to confront, then, is a gap in respect of the application of BFO to the security domain. The present report is, in part, an attempt to fill this gap, showing how BFO might be used as an upper-level ontology that supports the representation of security concepts as part of a wider effort to develop ontologies for the realm of IoT devices and Cyber-Physical Systems (CPSs). The following summarizes the main factors that informed our decision to rely on BFO for this purpose:[4]

- **Adoption:** BFO is one of the most widely used upper ontologies. In particular, BFO serves as the basis for both the OBO Foundry (Smith et al. 2007) and IOF (Smith et al. 2019; Drobnjakovic et al. 2022) initiatives. Widespread adoption provides opportunities for cross-domain integration, as well as exemplifying approaches to common modelling problems.

- **Standardization:** BFO has been designated an ISO standard[5] and BFO's ISO 21838-2 specification has been axiomatized in First-Order Logic, Web Ontology Language (OWL) 2, and Common Logic Interchange Format (CLIF) (Otte, Beverley, and Ruttenberg 2022).

- **Application to BIM:** IoT systems form part of the built environment, which establishes a link with work pertaining to BIM. In this respect, it is worth noting that BFO has been the focus of

[4] BFO is one of the ontologies surveyed by Partridge et al. (2020) as part of the effort to develop an Information Management Framework (IMF) for National Digital Twins (NDTs).
[5] ISO Standard No. 21838-2:2020. See https://www.iso.org/standard/74572.html.

recent research efforts within the BIM community (Park and Shin 2023; Tchouanguem et al. 2021).

- **Cyber Ontology:** Recent work has sought to apply BFO to the representation of computational artefacts and processes (Donohue et al. 2018). This serves as a valuable point of departure for ontological efforts that seek to model security-related concepts in the cyber-physical domain.[6]

- **4D Data Modelling:** BFO is intended to support both three- and four-dimensionalist approaches to data modelling. Arp et al. (2015, p. 124), for example, suggest that "BFO […] embrace[s] a four-dimensionalist perspective; but it combines this with a three-dimensionalist perspective for continuants, and does not attempt to reduce the one to the other."

3.2 Continuants

BFO adopts a view of reality according to which all things are regarded as **ENTITIES**.[7] **ENTITIES** are then decomposed into what are called **CONTINUANTS** and **OCCURRENTS**. In this sense, BFO assumes that all things—all entities—can be categorized as either continuants or occurrents (or a combination thereof). This distinction between continuants and occurrents is one that is common to many upper ontologies, and it reflects the basic metaphysical distinction between entities that persist through time, such as physical objects, and the occurrent entities (e.g., processes) in which these objects participate (see Rodrigues and Abel 2019). Arp et al. (2015) characterize this distinction as follows:

> **Continuants:** entities that continue or persist through time, including (1) independent objects (for example, things such as you and me); (2) dependent continuants, including qualities (such as your temperature and my height), and functions (such as the function of this switch to turn on this light); together with (3) the spatial regions these entities occupy at any given time […] (Arp, Smith, and Spear 2015, p. 87)
>
> **Occurrents:** entities that occur or happen, variously referred to as "events" or "processes" or "happenings," which we take to comprise not only (1) the processes that unfold in successive phases but also (2) the boundaries or thresholds at the beginnings or ends of such processes, as well as (3) the temporal and spatiotemporal regions in which these processes occur. (p. 87)

Figure 1 shows the various classes that comprise the continuant hierarchy in BFO. As can be seen from this figure, BFO distinguishes between three categories of continuant entity, namely, **INDEPENDENT CONTINUANTS**, **SPECIFICALLY DEPENDENT CONTINUANTS**, and **GENERICALLY DEPENDENT CONTINUANTS**. While it is not represented in the BFO continuant hierarchy, this labelling is suggestive of a more general distinction between independent continuants and dependent continuants, with the latter subsuming the categories of specifically and generically dependent continuants (Arp and Smith 2008). This distinction is clearly related to the notion of 'dependence'. In particular, independent continuants are defined as continuants that do not depend

[6] See https://opensource.ieee.org/cyber-ontology-working-group.
[7] See https://github.com/BFO-ontology.

on other entities for their existence. This contrasts with dependent continuants that do depend on other entities for their existence.

The form of dependence that is in play here is what is known (in philosophy) as *existential dependence*. The idea is that some things only exist in virtue of the existence of other things. My mass, for example, is specifically dependent on. If there is no me, then there is no mass that belongs to me (or that inheres in[8] me). My mass is a property of me, but it cannot exist independently of me. One might say that as long as I exist, then it must be the case that I also have a mass. That is true, but this is more a matter of metaphysical or nomological or (perhaps) natural necessity (see Toyoshima 2020) than it is one of existential dependence.[9] Note that I might have different masses at different points in time, and each of these masses are dependent on me, but it cannot be the case that a specific mass of me (e.g., 90kg) has an existence independent of me.

```
CONTINUANT
    SPECIFICALLY DEPENDENT CONTINUANT
        QUALITY
            RELATIONAL QUALITY
        REALIZABLE ENTITY
            DISPOSITION
                FUNCTION
            ROLE
    GENERICALLY DEPENDENT CONTINUANT
    INDEPENDENT CONTINUANT
        MATERIAL ENTITY
            OBJECT AGGREGATE
            FIAT OBJECT PART
            OBJECT
        IMMATERIAL ENTITY
            SITE
            CONTINUANT FIAT BOUNDARY
                ZERO-DIMENSIONAL CONTINUANT FIAT BOUNDARY
                ONE-DIMENSIONAL CONTINUANT FIAT BOUNDARY
                TWO-DIMENSIONAL CONTINUANT FIAT BOUNDARY
            SPATIAL REGION
                ZERO-DIMENSIONAL SPATIAL REGION
                ONE-DIMENSIONAL SPATIAL REGION
                TWO-DIMENSIONAL SPATIAL REGION
                THREE-DIMENSIONAL SPATIAL REGION
```

Figure 1. BFO continuants hierarchy. [Indentations reflect taxonomic (or sub-class of) relations.]

Arp et al. (2015) explicitly characterize the relation between dependent continuants and independent continuants as a form of existential dependence. They suggest that:

> A specifically dependent continuant is a continuant entity that depends on one or more specific independent continuants for its existence. Dependent continuants exhibit existential dependence in the sense that, in order for a dependent continuant to exist,

[8] The notion of inherence is a further form of philosophical relation. Typically, dependent continuants are said to inhere in the entities that bear these continuants. Such bearers are what BFO calls independent continuants.
[9] Metaphysical necessity is a philosophical notion that is related to, but not the same as, existential dependence.

> some other entity in which it inheres (intuitively, an entity enjoying a larger degree of concreteness) must exist also. (Arp, Smith, and Spear 2015, p. 95)

Earlier they note that the relation between dependent and independent continuants is one that:

> […] implies that the dependent entity is secondary (has diminished concreteness) in relation to the independent continuant that is its bearer. The latter is a three-dimensional thing that has material parts. The dependent entity, by contrast, has no material parts but is rather parasitic on the material thing that supports it. Material things cannot be parasitic on (or ontologically secondary to) other entities in this sense. (There is nothing more concrete than material things.) And from this it follows that an independent continuant, while it is an entity in which other entities (such as qualities) inhere, cannot itself inhere in anything. (p. 90)

The notion of existential dependence is best understood with respect to the relation between independent continuants and specifically dependent continuants. Independent continuants are most easily thought of as the physical (or material) objects that we see around us every day. They include the likes of a person, a person's nose, the chair a person is sitting on, the cat a person is stroking, the apple they are eating, and so on. Specifically dependent continuants are most easily thought of as the *properties* of these things. The shape of a person's nose, for example, is a property of the person's nose. Similarly, the colour of an apple is a property of the apple. These properties are said to ***inhere in*** entities that bear these properties. Such entities are what BFO calls independent continuants.

As can be seen from Figure 1, BFO distinguishes between two types of specifically dependent continuant. These are: **QUALITIES** and **REALIZABLE ENTITIES**. Qualities comprise the overt or manifest properties (or attributes or features) of an independent continuant. They include things such as mass, length, shape, and colour. Unlike realizable entities, qualities are properties that are observable in the here-and-now; they do not depend on processes for their manifestation. This contrasts with realizable entities, which are deemed to rely on processes for their manifestation (Arp and Smith 2008).

The distinction between qualities and realizable entities reflects the philosophical distinction between what are called categorical and dispositional properties.[10] In philosophy, a categorical property is a characteristic or attribute or feature that is inherent or essential to an entity, independent of any other factors or conditions. It is a property that is not contingent upon anything else but is rather a fundamental feature of the thing itself. All qualities are categorical properties. The shape of my nose, for example, is a categorical property (quality) of my nose. Similarly, the mass of my body is a categorical property (quality) of my body. Neither of these properties are dependent on something else. This contrasts with dispositional properties, which are the sorts of properties that rely on something else for their manifestation, actualization, or realization. My ability to speak English, for example, is not something that is a manifest property of me; instead, my ability to speak English is manifest in my actually speaking English. In this sense, my ability to speak English is not immediately manifest in me; it is only manifest when I participate in an act of speaking, where an act

[10] As noted by Goldfain et al. (2010, p. 401): "BFO embraces a distinction between categorical properties (e.g., triangularity) and dispositional properties (e.g., fragility). BFO makes this distinction by partitioning specifically dependent continuants (i.e., individual entities that depend for their existence on a specific bearer) into qualities (categorical properties) and realizable entities (including dispositional properties and roles)."

of speaking is a process (or, more generally, an occurrent). Accordingly, my ability to speak English cannot be independent of other entities (in regard to its manifestation) for it is only manifest when we encounter another sort of entity, namely, a (speaking) process. In BFO parlance, an ability is a type of **DISPOSITION**,[11] which is a type of **REALIZABLE ENTITY**. Such entities are deemed to be ***realized in*** **OCCURRENTS**, where the notion of an **OCCURRENT** subsumes things like processes and events.

The following definitions highlight the distinction between the various types of specifically dependent continuant in BFO:

> **QUALITY** =*def.* A specifically dependent continuant that is exhibited if it inheres in an entity or entities at all (a categorical property) (Goldfain, Smith, and Cowell 2010).
>
> **REALIZABLE ENTITY** = *def.* A specifically dependent continuant that inheres in independent continuant entities and is not exhibited in full at every time in which it inheres in an entity or group of entities (Goldfain, Smith, and Cowell 2010).
>
> **DISPOSITION** = *def.* A disposition is a realizable entity which is such that, if it ceases to exist, then its bearer is physically changed, and whose realization occurs in virtue of the bearer's physical make-up when this bearer is in some special circumstances (Goldfain, Smith, and Cowell 2010).
>
> **ROLE** = *def.* A role is a realizable entity that exists because there is some single bearer that is in some special physical, social, or institutional set of circumstances in which this bearer does not have to be, and the realizable entity is not such that, if it ceases to exist, then the physical make-up of the bearer is thereby changed.

Note that while **ROLES** and **DISPOSITIONS** are both types of **REALIZABLE ENTITY**, they are not the same. At first sight, this might seem a little odd, for roles have much in common with dispositions. Someone's role as a cybersecurity analyst, for example, is not something that is manifest in a person per se; it is more something that is realized in the performance of certain tasks, specifically, those we associate with the occupational role of a cybersecurity analyst. According to the proponents of BFO, however, there is an important distinction between **ROLES** and **DISPOSITIONS**. This turns on what is referred to as the *internal grounding assumption* (see Goldfain et al. 2010).[12] According to this assumption, dispositions are grounded in categorical properties (or qualities) that inhere in their bearers. Consider the fragility of a vase (see Figure 2). In BFO, a vase is an independent continuant, while the fragility of the vase is a disposition. The fragile disposition is manifest in a certain set of circumstances. For example, when the vase is dropped on a hard surface, then its fragility is manifest in a breaking process. This is what makes fragility a dispositional property—a property that is actualized or manifest in a certain sort of process. At the same time, however, it is not the case that the fragility of the vase is independent of the manifest features (the qualities) of the vase. The vase is fragile due to the fact that it is made of a material with a certain molecular configuration, and it is this molecular configuration that makes the vase fragile. For this reason, we say that the relevant disposition (the fragility) is *grounded in* the qualities of the vase. If these qualities were changed, then

[11] It should be noted that not everyone accepts the idea that abilities ought to be cast as dispositional properties (see Chemero 2009).
[12] McKitrick (2018, chap. 8) refers to this as the *intrinsic disposition thesis*.

the vase might cease to be fragile, but is it hard to imagine a set of circumstances in which the vase would no longer be fragile in the absence of some sort of change to the material from which it was made.

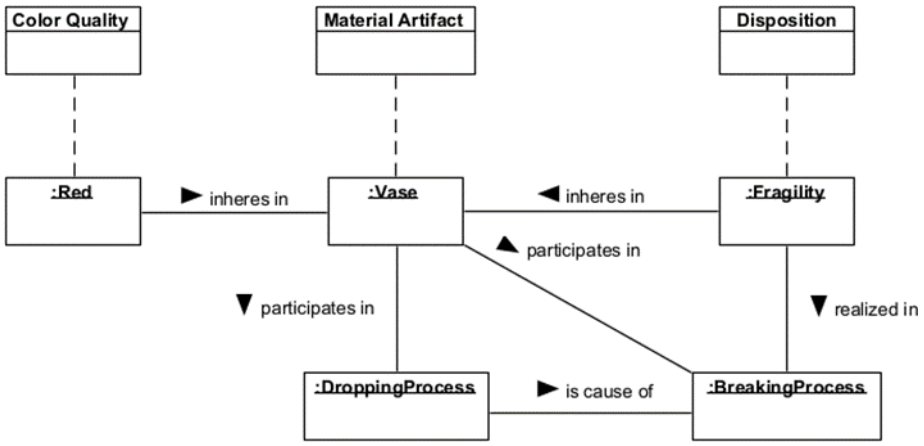

Figure 2. The fragility of a (red) vase is realized in a breaking process. [Dashed lines symbolize instantiation relations.]

While dispositions are deemed to be internally grounded in the properties (the qualities) of their bearers, roles are not so grounded. Someone's role as a cybersecurity analyst, for example, is not something that supervenes on a more fundamental set of categorical properties that are intrinsic to the bearer of the role. Someone could lose their job as a cybersecurity analyst without undergoing a corresponding shift in their intrinsic properties. Similarly, someone could come to occupy a certain role without undergoing any form of constitutional change. To be sure, there may be properties of a thing that make it suitable for the occupation of certain roles, but these things are not necessarily lost or gained as the result of the occupation of these roles.

In addition to independent continuants and specifically dependent continuants, BFO recognizes a third type of continuant known as a generically dependent continuant. Like specifically dependent continuants, generically dependent continuants are deemed to be continuants that depend on other continuants, namely, independent continuants. Unlike specifically dependent continuants, however, generically dependent continuants are able to be copied or transferred to other entities. For the most part, generically dependent continuants are used to support the modelling of information artefacts and the flow of information between these artefacts. For this reason, we postpone a discussion of these continuants to Section 3.6, where we discuss issues of information modelling.

```
OCCURRENT
    SPATIOTEMPORAL REGION
    PROCESS BOUNDARY
    TEMPORAL REGION
        ZERO-DIMENSIONAL TEMPORAL REGION
        ONE-DIMENSIONAL TEMPORAL REGION
    PROCESS
        HISTORY
        PROCESS PROFILE
```

Figure 3. BFO occurrents hierarchy.

3.3 Occurrents

Figure 3 shows the occurrent portion of the BFO class hierarchy. In BFO, the main category of occurrents are processes, which are things like the process of sleeping, the process of meiosis, an ageing process, the course of a disease, the beating of a heart, or the process of assessing risk. In philosophy, the metaphysical category of occurrents is deemed to consist of events, states, and processes (see Kaiser and Krickel 2017; Rodrigues and Abel 2019). In BFO, however, there are no classes to represent states and events. This can be a little confusing, for other ontologies, including those developed in the security domain, tend to talk of events rather than processes (Sales et al. 2018; Oliveira et al. 2022). The extent to which BFO recognizes a genuine distinction between events and processes remains unclear. Quite plausibly, events are just another type of process, or they are perhaps understood as occurrents that have no constituent temporal parts. They are, perhaps, simply occurrents that happen in an instantaneous fashion, which is to say they are occurrents without a discernible temporal duration: they are simply occurrents that ***occur on*** **TEMPORAL REGIONS** that lack a temporal extent (i.e., events ***occur on*** time instants—or, in BFO parlance, **ZERO-DIMENSIONAL TEMPORAL REGIONS**).

The extent to which BFO recognizes states as occurrent entities is similarly obscure. As noted above, BFO does not include a class to explicitly represent states. The CCO does, however, include a STASIS class, which is defined as follows:

> **STASIS** =*def.* A Process in which one or more Independent Continuants endure in an unchanging condition.

The various subtypes of this class (see Figure 4) seem to suggest that states are being understood as a specialized type of process. In support of this interpretation, an introductory overview of the CCO describes the **STASIS** class as follows:

> Note that although most processes involve an object actively changing something or passively undergoing change, PROCESS also includes object states, in which an object does not change with respect to one of its attributes over some period of time. Thus, we can describe a person (an object) having the role of surgeon (an attribute) over some specific period of time. This notion of an object state is captured by the CCO class STASIS […] (CUBRC 2020a, p. 7)

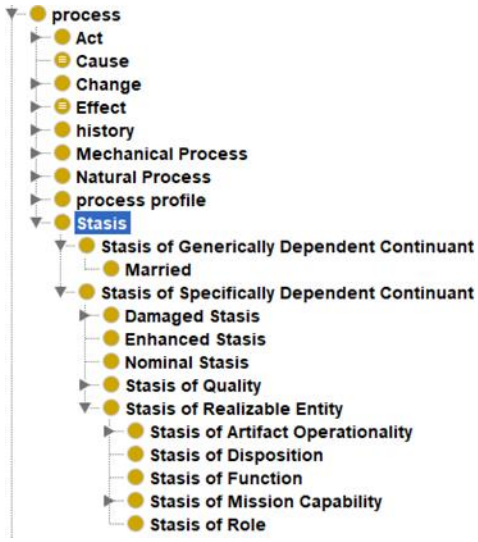

Figure 4. The CCO relies on the STASIS class to represent the states of things.

In addition to the **STASIS** class, the CCO includes a class labelled **CHANGE**, which is also a subclass of **PROCESS** (see Figure 4). This class can be used to represent information about changes to object properties, such as when an object gains or loses a particular role. In Figure 5, for example, we show how one can represent the gain and loss of the U.S. presidential role.

In BFO, the attributes, properties, or features of processes are represented using the **PROCESS PROFILE** class. In particular, a process profile is characterized as:

> […] an abstraction of some relevant facet of a process (typically, a change or rate of change of some object attribute). For example, the speed of some vessel (the rate of its distance travelled divided by the time elapsed) can be represented as a process profile of the movement in which that vessel participates. (CUBRC 2020a, p. 8)

Figure 6 exemplifies the use of process profiles to capture information about the speed of an independent continuant (in this case, an aerial drone) that participates in an act of motion (i.e., a type of process). As can be seen from this figure, speed information is represented as an instance of the **PROCESS PROFILE** class that forms part of an instance that is of type **ACT OF MOTION**. Note that **PROCESS PROFILES** are, themselves, types of **PROCESSES** that are connected to the relevant process instance via a ***has process part*** relation. In short, the relationship between **PROCESSES** and **PROCESS PROFILES** is one of processual parthood.

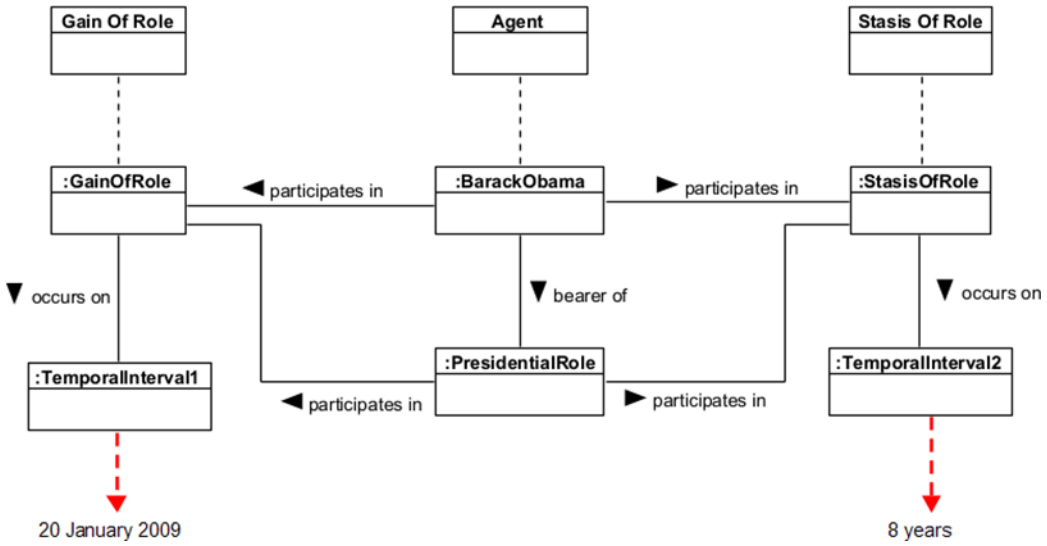

Figure 5. The CCO includes classes to represent changes to the properties of things (in this case, roles).

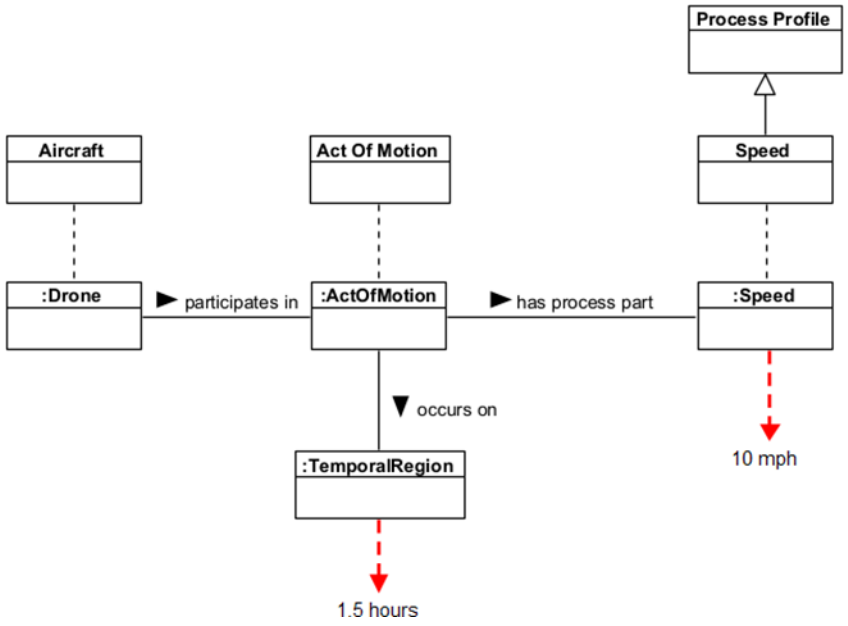

Figure 6. The use of process profiles to represent the speed of an unmanned aerial drone.

In BFO, every independent continuant has a **HISTORY**, which is a **PROCESS**. In BFO, histories are defined as follows:

> **HISTORY** $=_{def.}$ A history is a process that is the sum of the totality of processes taking place in the spatiotemporal region occupied by a material entity or site, including processes on the surface of the entity or within the cavities to which it serves as host.

As should be clear from this definition, histories provide a way of referring to all of the processes in which a given object (independent continuant) participates. In short, all the processes in which an

independent continuant participates form part of its history. Histories, themselves, occupy spatiotemporal regions, which are defined as an occurrent entity that is part of spacetime. The **HISTORY** class establishes a point of contact with so-called 4D ontologies or 4D approaches to data modelling, which assume that all entities can be seen as having both a spatial and temporal extension (see West 2011).

3.4 Common Core Ontologies

As an upper-level ontology, BFO does not provide support for the representation of domain-specific entities; it simply provides an abstract framework in which domain-specific entities can be represented. The application of BFO to more specialized domains is supported by the CCO,[13] which is a suite of twelve ontologies serving as a mid-level extension to BFO (see Figure 7). These ontologies are as follows:

- **Extended Relation Ontology:** This ontology is designed to represent many of the relations (i.e. object properties) that hold between entities at the level of the mid-level CCOs.
- **Modal Relation Ontology:** This ontology contains modal counterparts to the relations contained in the CCO extended relation ontology.
- **Geospatial Ontology:** This ontology is designed to represent sites, spatial regions, and other entities, especially those that are located near the surface of Earth, as well as the relations that hold between them.
- **Time Ontology:** This ontology is designed to represent temporal regions and the relations that hold between them.
- **Information Entity Ontology:** This ontology is designed to represent generic types of information as well as the relationships between information and other entities.
- **Agent Ontology:** This ontology is designed to represent agents, especially persons and organizations, and their roles.
- **Artifact Ontology:** This ontology is designed to represent artefacts that are common to multiple domains along with their models, specifications, and functions.
- **Facility Ontology:** This ontology is designed to represent buildings that are designed to serve some specific purpose, and which are common to multiple domains.
- **Currency Unit Ontology:** This ontology is designed to represent currencies that are issued and used by countries
- **Event Ontology:** This ontology is designed to represent processual entities, especially those performed by agents, that occur within multiple domains.
- **Quality Ontology:** This ontology is designed to represent a range of attributes of entities especially qualities, realizable entities, and process profiles.

[13] See https://github.com/CommonCoreOntology/CommonCoreOntologies.

- **Units of Measure Ontology:** This ontology is designed to represent standard measurement units that are used when measuring various attributes of entities.

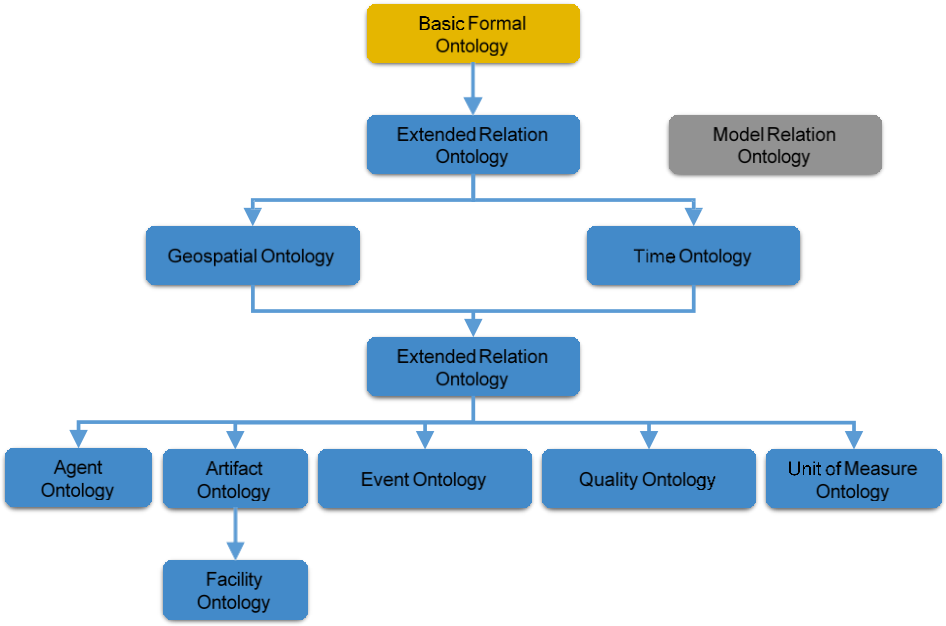

Figure 7. The Common Core Ontologies suite. [Arrows reflect the import structure of the CCO suite.]

Note that while these ontologies feature many hundreds of classes, they are all built on the back of BFO. That is to say, all the classes within the CCO are represented as subtypes of the classes contained in BFO. Also note that these ontologies are intended as a *mid-level* extension to BFO. The CCO is not intended to support the development of ontologies in any particular domain; rather, they are intended to be across many different domains. In this sense, their scope is less constrained than would be the case with a domain-level ontology.

An interesting, albeit important, feature of the CCO is the inclusion of a modal relation ontology. As noted above, the modal relation ontology is an ontology that serves as a counterpart to the extended relation ontology. In fact, all the relations in the extended relation ontology are included in the modal relation ontology. The only distinction between these ontologies relates to the use of a particular namespace. While relations in the extended relation ontology rely on the `cco:` namespace, those in the modal relation ontology rely on the `mro:` namespace.

The relations in the modal relation ontology are intended to serve as modal counterparts to those contained in the extended relation ontology. What this means is that one can use modal relations to represent counterfactual states-of-affairs, such as how things could be (or perhaps should be) in the future (see Jensen et al. 2018). This is important, for the realist orientation of BFO makes it difficult to talk about non-existent states-of-affairs. By definition, things that are non-existent, do not exist, and BFO is concerned with things that do exist, or that are, in some sense, 'real' (recall the commitment of BFO to ontological realism). Despite this, it is widely recognized that there is a need to talk about counterfactual states-of-affairs. This is particularly important when it comes to issues of risk and security, for such domains rely on a capacity to predict how things might evolve into the future. The notion of risk, for example, doesn't appear to make much sense if we are unable to refer

to things that *might* happen in the future. A similar point is made by Sales et al. (2018) as part of their attempt to provide an ontology of risk:

> The classical view of events assumes that they are immutable entities and that only past events truly exist as genuine perdurants (occurrences). However, accounting for future events (which is the case for envisioned experiences) seems to be unavoidable for any theory of risk, as uncertainty and possibility are core aspects of this concept. This means that we need to refer to future events—whose expected temporal properties are not completely fixed—as first-class citizens in our domain of discourse. As bold as this assumption may seem […] conceptualizing risk with no reference to the future would sound as an oxymoron to us. So, we shall talk of expected events as regular entities of our domain, not differently from, say, a planned air trip in a flight reservation system. (Sales et al. 2018, pp. 128–129)

In BFO, the traditional approach to modality has been one of reliance on realizable entities, specifically dispositions. Goldfain et al. (2010), for example, note that:

> As part of its realist orientation, BFO attempts to avoid treatments of modality (necessity, possibility) in terms of special entities such as possible worlds in favor of a focus on objects existing in the present, actual world. Dispositions provide a formal mechanism for taking account of future manifestations (BFO occurrents) in terms of what is true of the underlying independent continuants in the present; roughly, dispositions say how something is in terms of what it has the built-in potential to do or suffer. (Goldfain, Smith, and Cowell 2010, p. 142)

The inclusion of modal relations in the CCO is driven by the need to represent situations where there exists some sort of representation of a counterfactual state-of-affairs, without that representation qualifying as a disposition. Consider, for example, the expectation that it will rain tomorrow. This expectation is clearly about a future state-of-affairs, specifically, the nature of the weather tomorrow, and the expectation is (let's assume) encoded in some material entity. The details of this encoding need not concern us here. The encoding could be something like a mental representation that exists in someone's head (a belief), or it could be a collection of written words and other material symbols (see Clark, 2006) that form part of a meteorological forecast. For present purposes, what matters is that there is something that is referring to a state-of-affairs that does not exist at the present point in time. Modal relations provide us with a means of referring to this potential (although not yet actual) state-of-affairs. Rather than stating that it is currently raining, or that it has rained at some particular point in the past, we rely on modal relations to represent the fact that the expectation pertains to some future state-of-affairs, one that may or may not come to pass.

Note that this is not the same as saying that there is some atmospheric entity that is disposed to rain tomorrow. This may, of course, be true, in the sense that there is a genuine disposition that exists in the relevant region of the atmosphere that will be realized in a raining process at some point in the future. There is, however, no guarantee that the atmosphere is the genuine bearer of this disposition; all we have is the belief or expectation that it might possibly rain, not that it will actually rain.

Another common use of modal relations relates to situations where we want to express normative constraints. Perhaps, for example, we want to specify that a sequence of actions *should* unfold according to some plan. The problem, here, is that not everything goes according to plan. Regardless

of how meticulous we are with regard to the specification of a plan, there is no guarantee that things will evolve as we expect them to. As the heavyweight boxer, Mike Tyson, once quipped, "Everyone has a plan until they get punched in the mouth." When it comes to the realm of ontologies, we often want to represent the fact that something should have unfolded in a certain way, even if it does not conform to that expectation in reality. In this situation, we confront a mixture of entities, some of which are real and some of which are not real. The things that are real include the plan and its constituent expectations pertaining to the sequence of actions that ought to be performed. Also lying within the realms of the real are the actual actions that were performed as part of an attempt to follow the plan. What is not real, however, are the things that are referred to by the expectations contained within the plan. These things are possible states-of-affairs that may or may not come to pass. Such things do not exist at the time the plan was created, but nor do they exist when the plan is implemented. This may sound a little odd but bear in mind that the actual actions may not conform to the plan, so there is no sense in which the mere description of future actions refers to the actions that actually exist.

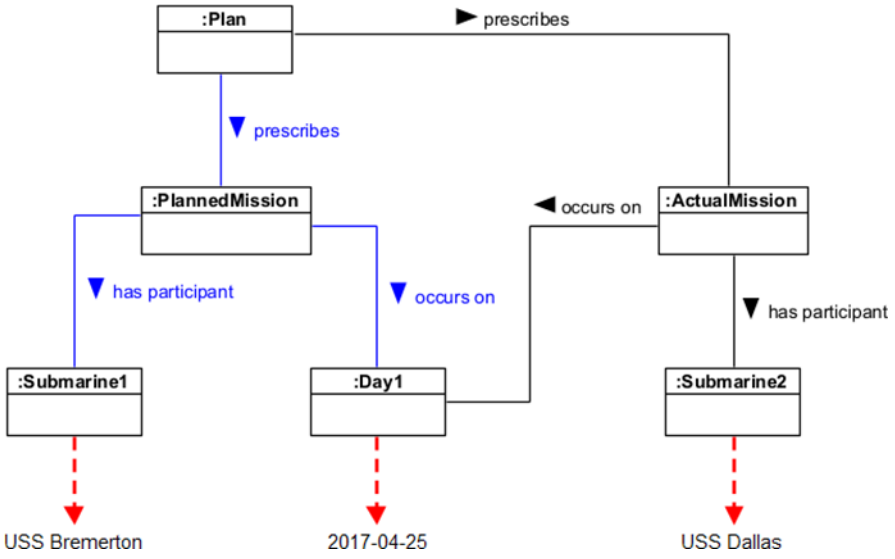

Figure 8. Modelling actual and planned missions. [Modal relations are symbolized in blue font.]

Figure 8 exemplifies the use of modal relations in a simple planning scenario. To distinguish between standard and modal relations, we render modal relations in blue font (the same approach is adopted throughout the remainder of the present report). In this case, we have a `:PLAN` that prescribes a course of action (i.e., `:PLANNEDMISSION`). The details of the plan are represented via modal relations. Thus, the plan prescribes that a particular mission should occur on a certain day and that it should involve a submarine called the *USS Bremerton*. As it happens, the actual mission (denoted by `:ACTUALMISSION`) occurred on the same day as that given in the plan, but it involved a different submarine, namely, the *USS Dallas*. The use of modal relations thus enables us to specify what should have occurred, which provides the basis for comparisons as to what actually occurred.

3.5 The Common Core Cyber Ontology

In recent years, a number of domain-level ontologies have been built on top of the CCO. One such ontology is of particular interest to the present research effort, for it seeks to extend the CCO to the

realm of cyber objects and cyber processes (collectively: cyber entities). The ontology—called the C3O—is intended to support the representation of entities relevant to the digitization, manipulation, and transfer of information using telecommunication networks, especially as they pertain to activities in cyberspace (Donohue et al. 2018).[14] Some of the entities included within the C3O are information processing artefacts and their functions, networks and their components, software, protocols and standards, users and their permissions, data transformation and encoding, and cyber-attacks and their objectives. This, it should be clear, establishes an important point of contact with the SOfIoTS project, although it should also be noted that the C3O does not include specific support for IoT devices or CPSs. The other thing to note is that the C3O remains a work in progress. Recent work has sought to extend the C3O with the entities described by the MITRE ATT&CK® Matrices for Enterprise IT and for Mobile devices, culminating in the MITRE ATT&CK Matrix Ontology (MAMO) ontology[15] (see also Naray, Haugh, and Wartik 2022). We will not attempt to survey the details of MAMO in the present report. Suffice to say, MAMO includes classes to represent cyber-attacks, cyber objectives, cyber agents, and acts of cyber threat mitigation. Some of the types of cyber-attacks recognized by MAMO are depicted in Figure 9.

Figure 9. Types of cyber-attack included in MAMO.

3.6 Information Modelling

In CCO, support for information modelling is provided via the information entity ontology. This ontology includes many of the terms that were defined in the Information Artifact Ontology (IAO),

[14] See https://opensource.ieee.org/cyber-ontology-working-group.
[15] See https://opensource.ieee.org/cyber-ontology-working-group/c30-extension.

which serves as the precursor to the information entity ontology (Smith et al. 2013; Smith and Ceusters 2015). Both these ontologies are concerned with the representation of what might be called information entities or information artefacts. The use of these entities forms that basis of BFO's approach to information and data modelling.

BFO adopts a particular approach to the modelling of information entities. At the core of this approach is a tri-fold distinction between information content, the objects that serve as the physical carriers (or vehicles) of information content, and the patterns that encode information content. This distinction is best understood with respect to a particular example. Consider a conventional (paper-based) book. The book, itself, is a physical object, or material entity. More specifically, it is an object that bears (or carries) information pertaining to a particular topic. In CCO, this object is what is known as an **INFORMATION BEARING ENTITY**, which is a type of **MATERIAL ENTITY**. Information bearing entities are defined as follows:

> **INFORMATION BEARING ENTITY** =$_{def.}$ An information bearing entity is a material entity that has been created to serve as a bearer of information. (Smith et al. 2013).[16]

A book is thus an information bearing entity, which serves as the carrier (or bearer of) information. The information, itself, is represented via another entity, called an **INFORMATION CONTENT ENTITY**. In CCO, **INFORMATION CONTENT ENTITIES** are cast as **GENERICALLY DEPENDENT CONTINUANTS**. Like **SPECIFICALLY DEPENDENT CONTINUANTS**, **GENERICALLY DEPENDENT CONTINUANTS** are a type of dependent continuant. That is to say, they depend for their existence on some other entity. Unlike **SPECIFICALLY DEPENDENT CONTINUANTS**, however, **GENERICALLY DEPENDENT CONTINUANTS** may be borne by different entities at different times, or they may be borne by multiple entities at the *same* time. Returning to the book example, a single book is the carrier of a certain body of information content, but this information content may be borne by other books. In particular, the content of one book may be precisely duplicated by another book. In this case, we have two physical books (two **INFORMATION BEARING ENTITIES**), but both these books carry the *same* information content. Such information content will exist as long as there is at least one book that acts as its carrier. Suppose that the two aforementioned books were the only two books in existence. One of these books could be destroyed, but this would not entail a corresponding destruction of the information content carried by that book. This would only be the case if *both* books were destroyed, and there were no other physical carriers of the relevant information content. As long as one book remains in existence, then the information content will be preserved. What is more, the longevity of the information content can be preserved if it is transferred to other media. If the information content of a book is copied into, let's say, a digital format, then the information content will be borne by additional physical objects, specifically, digital objects. At this point, the last remaining copy of a physical book could be destroyed, but as long as there remains at least one digital object that is the carrier of the book's information content, then the information content of the book will remain in existence.

What we see here, then, is a basic distinction between the physical object that bears some information content and the information content that is borne by that physical object. In CCO, the physical object is represented by the aforementioned **INFORMATION BEARING ENTITY**, while the content is

[16] There is, of course, a sense in which a book is specifically *designed* to act as the carrier of information content. This establishes the status of the book as an **INFORMATION BEARING ARTIFACT**, which is a particular type of **INFORMATION BEARING ENTITY**.

represented by an **INFORMATION CONTENT ENTITY**. **INFORMATION CONTENT ENTITIES** are then defined as follows:

> **INFORMATION CONTENT ENTITY** =*def.* Information content entities are about something in reality (they have this something as a subject; they represent, or mention or describe this something; they inform us about this something) (Smith et al. 2013).

As suggested by this definition, **INFORMATION CONTENT ENTITIES** are typically about something—they have as their referent some other entity, which may or may not be a material entity. In CCO, the link between an information content entity and the thing the information content entity is about is established via the ***is about*** relation. The range of this relation can refer to anything that qualifies as an **ENTITY**. A single information content entity could thus be about a material entity, such as physical (paper-based) book (i.e., an **INDEPENDENT CONTINUANT**, or it could be about one of the properties of that book, such as a particular quality of the book (i.e., a **SPECIFICALLY INDEPENDENT CONTINUANT**). What is more, there is nothing to prevent one information content entity from being about another information content entity. That is to say, there is no reason why one information content entity cannot refer to another entity that qualifies as a **GENERICALLY DEPENDENT CONTINUANT**. This is particularly useful when it comes to representing provenance-related information, such as information about when (and how) other information was generated. Finally, information content entities can refer to occurrent or processual entities, such as **PROCESSES**. This enables us to represent information about such processes, such as measurements of the attributes (or **PROCESS PROFILES**) associated with particular processes.

While information content and information bearing entities are the two main classes of information entities within CCO, there is a further entity to consider. Such entities are called **INFORMATION QUALITY ENTITIES**. Information quality entities provide a bridge between information content entities and information bearing entities. In particular, information content entities are said to be concretized in information quality entities, which are then borne by information bearing entities. From an ontological standpoint, information quality entities are a type of quality, which puts them in the category of specifically dependent continuants. From a more common-sense perspective, information quality entities can be thought of as the physical patterns that encode information content. Consider, again, the case of a book. Each page within the book is the bearer of a series of symbolic tokens (e.g., words) that are organized in a particular pattern. In CCO, this pattern is a **QUALITY** that is deemed to 'concretize' the information content that is carried by the physical object (the book) in which the pattern inheres. It is this notion of concretization that serves as the basis for the definition of information quality entities:

> **INFORMATION QUALITY ENTITY** =*def.* An information quality entity is a quality that is the concretization of some information content entity (Smith et al. 2013).

It is important to note that the same information content entity may be concretized in different ways depending on the nature of the information bearing entity that bears the information content. Consider the case where the information content of a physical book is copied into a digital format. In this case, we have a single information content entity that (generically) depends on two information bearing entities (i.e., a book and a digital object), but the information qualities that concretize (or encode) the information content are not the same. In the case of a book, the information is concretized by patterns of ink, while in the case of the digital object, the information is concretized

by a pattern of 1s and 0s in the magnetic coating of a computer's hard drive. Here, then, we have two information bearing entities and two information quality entities, but all these entities are connected to a single information content entity. As noted by Otte et al. (2022):

> Because information content entity is a direct subclass of generically dependent continuant, an information content entity may generically depend on one or more material entities. One example is the content of a novel may be concretized by patterns of ink in multiple physical books or may be concretized by the digital patterns in different network servers; when this occurs, the novel (an information content entity) then generically depends on the physical books and network servers. (Otte, Beverley, and Ruttenberg 2022, p. 8)

While information quality entities are represented as a particular type of **QUALITY** entity, this does not mean they cannot serve as the basis for other entities. As noted in Section 3.2, from a philosophical standpoint, qualities can be understood as categorical properties, but categorical properties can provide the basis for non-categorical properties, specifically, dispositional properties. In this sense, then, it is possible to imagine situations in which the instantiation of an information quality entity (a **QUALITY**) entails the instantiation of a dispositional property (or a **DISPOSITION**). This point is highlighted by Smith and Ceusters (2015):

> All concretizations are qualities in the BFO framework. Such qualities can serve as the basis for dispositions. When we concretize a lab test order by reading the text of the order on our screen, then in addition to the mental quality that is formed in our mind as we read the text, there is also a disposition to be realized in our actions of carrying out the relevant test. This disposition may come into being simultaneously with the mental quality created through our understanding of the text, but it is still dependent on this quality, as is shown by the fact that the latter may exist even in the absence of any accompanying disposition. (Smith and Ceusters 2015, p. 1)

Perhaps the best way of understanding this relationship between information qualities and dispositions is via the notion of a mental (and or cognitive) representation. According to the proponents of BFO, mental representations are a particular type of quality, namely a mental quality (Limbaugh et al. 2020, 2020; Smith and Ceusters 2015). These terms are defined as follows:

> **MENTAL QUALITY** =$_{def.}$ A mental quality is a quality that specifically depends on an anatomical structure in the cognitive system of an organism.
>
> **MENTAL REPRESENTATION** =$_{def.}$ A representation that is a mental quality.

Mental representations are typically thought of as things that are about something. From a BFO standpoint, these representations are patterns of neural activity that inhere in a person's brain, where the brain is an anatomical structure, which is a material entity and thus an independent continuant. In this scenario, then, the brain is an information bearing entity, the pattern of neural activity is an information quality entity, and the thing that is encoded by the pattern of neural activity is an information content entity. The content of the mental representation—the thing the mental representation is about—is then captured via the ***is about*** relation. This helps us understand what is going on in the lab-test-scenario discussed by Smith and Ceusters (2015) in the above quotation. At the outset, we have an information content entity that is concretized by a particular pattern of text

that appears on a screen. By reading this text, the reader acquires a mental representation whose content is (one hopes) the same as that encoded by the text. So, it is by reading the on-screen text that a person acquires a mental representation whose content is the same as that encoded by the text. The instantiation of this representation then provides the basis for the instantiation of a disposition pertaining to the performance of certain actions.

Figure 10 summarizes the relationships between the three types of information entity discussed above. As can be seen from this figure, CCO adopts a particular approach to the representation of literal values, such as numeric, string, and date/time values. [In Figure 10, the **INFORMATION BEARING ENTITY** class is shown to have a particular type of literal value, namely a decimal value. In practice, however, any type of literal value can be associated with an information bearing entity.] It is thus the information bearing entity—the carrier entity—that is attached to particular literal values, not the information content entity. This results in the following general pattern for information modelling:

INFORMATION CONTENT ENTITY ***is about*** **ENTITY**

INFORMATION BEARING ENTITY ***is carrier of*** **INFORMATION CONTENT ENTITY**

INFORMATION CONTENT ENTITY ***is concretized as*** **INFORMATION QUALITY ENTITY**

INFORMATION QUALITY ENTITY ***inheres in*** **INFORMATION BEARING ENTITY**

INFORMATION BEARING ENTITY ***has value*** *Literal*

Given that this pattern entails the creation of multiple objects (e.g., OWL individuals), CCO includes the ***is tokenized by*** relation, which is used to link literal values to instances of **INFORMATION CONTENT ENTITY** (see CUBRC 2020b). This reduces the aforementioned (triple) pattern to a single statement, namely:[17]

INFORMATION CONTENT ENTITY ***is tokenized by*** *Literal*

As noted above, an information content entity is linked to the entity that it is about by means of the ***is about*** relation. This relation has four sub-properties, each of which corresponds to a different form of 'aboutness'. These relations are as follows:

> ***describes*** $=_{def.}$ Information that describes some entity, such as the content of a report that describes an accident.
>
> ***designates*** $=_{def.}$ Information that designates some entity, such as an identifier.
>
> ***prescribes*** $=_{def.}$ Information that prescribes an entity, such as the content of a plan that prescribes a given sequence of actions.
>
> ***represents*** $=_{def.}$ Information that represents some entity, such as a photograph image that represents a particular object.

[17] In the OWL serialization of CCO, the ***is tokenized by*** relation is defined as an annotation property.

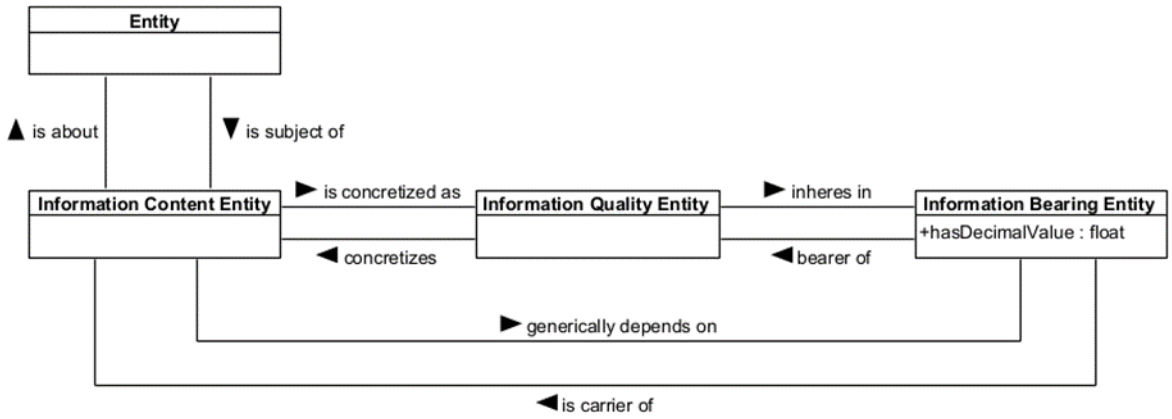

Figure 10. Relationship between information entities.

The modal variants of these relations (e.g., ***mro:prescribes***) can be used to refer to entities that have no concrete existence in the here-and-now. For example, a design specification could refer to the features of a device that is yet to be implemented. In this case, the design specification would exist as an information content entity that generically depends on one or more information bearing entities (e.g., technical documents). However, the thing the design specification is about would not exist until the design specification had been translated into a physical device. Figure 11 shows how this state-of-affairs is represented in CCO. Here, the design specification is represented by the object labelled `:DESIGNSPECIFICATION`, which is an **`INFORMATION CONTENT ENTITY`**. This specification mandates that the relevant artefact (`:IOTDEVICE`) needs to participate in certain processes (specifically, **`INFORMATION TRANSFER PROCESSES`**) with a certain speed. In particular, it ought to be designed in such a way that it is able to communicate information at a rate of 42 megabits per second. The `:DESIGNSPECIFICATION` also includes information about the properties of the human users who will ultimately use the device in the context of some process (denoted by `:PROCESS`). This information is represented via the `:USERSPECIFICATION` and `:CAPABILITYSPECIFICATION` objects. The first of these objects specifies that a user will use the target device to perform a process, and the second object specifies the capabilities that need to inhere in the user in order for them to use the device in the manner suggested by the design specification. While all this information is contained in the design specification, there is no sense in which any of the information ***`is about`*** entities that exist at the time the `:DESIGNSPECIFICATION` is produced. For this reason, most of the relationships depicted in Figure 11 are the modal counterparts of the standard relations included in the extended relation ontology (see Section 3.4).

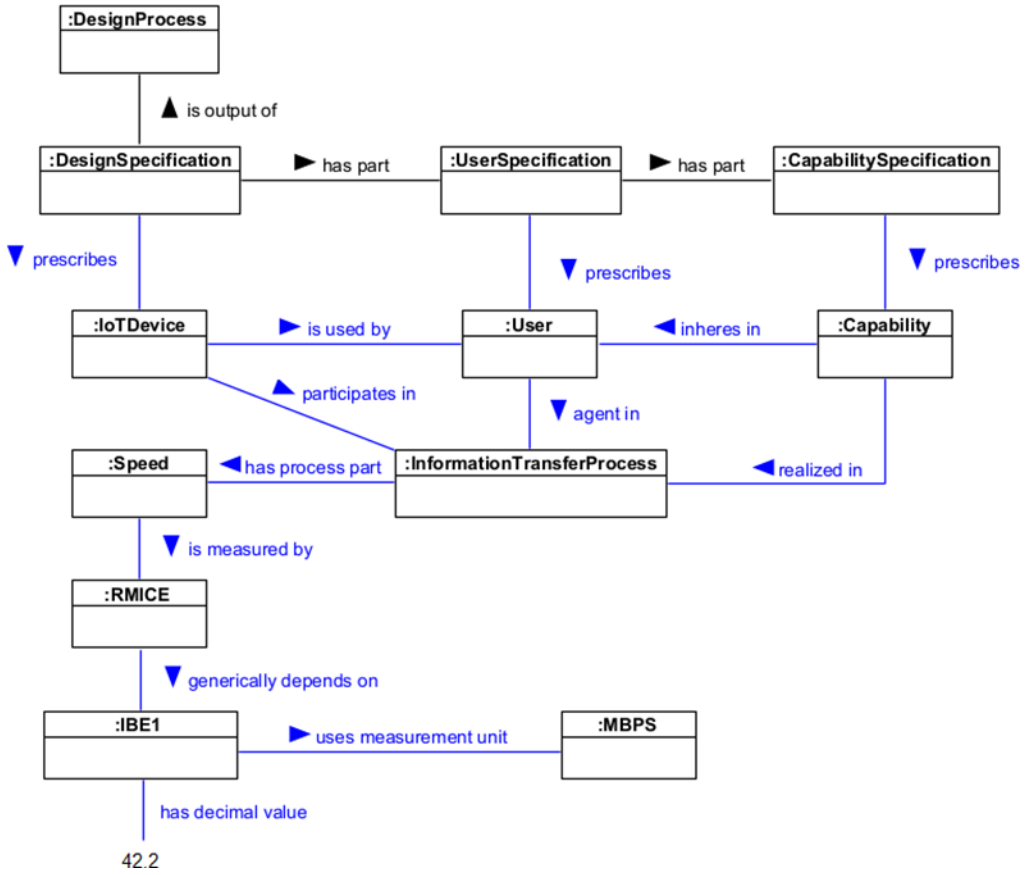

Figure 11. Design specification for an IoT device. [Blue font symbolizes modal relations. Acronyms: RMICE (Ratio Measurement Information Content Entity), IBE (Information Bearing Entity), MBPS (MegaBits Per Second).]

4 SECURITY CONCEPTS IN BASIC FORMAL ONTOLOGY

In this section, we discuss how common security concepts, such as the concepts of risk, threat, security goal, vulnerability, and so on might be accommodated by a BFO-conformant security ontology. Some common security concepts are presented in Section 4.1. In subsequent sections, we provide a brief analysis of each of these concepts and discuss how it might be represented using terms drawn from the ontologies discussed in Section 3, specifically, BFO, CCO, and the C3O.

4.1 Common Security Concepts

While there are significant differences between security ontologies, there is at least a degree of consensus regarding the concepts that ought to be included in a security ontology. According to ChatGPT, the concepts most commonly associated with security modelling are as follows:

- **Threats:** Threats refer to potential events or actions that can exploit vulnerabilities and cause harm to a system or organization's security. Identifying and understanding threats is essential for effective security modelling.
- **Vulnerabilities:** Vulnerabilities are weaknesses or gaps in a system's security controls that can be exploited by threats. Analyzing vulnerabilities helps identify areas that require protection or mitigation measures.
- **Risk:** Risk is the potential for loss or harm resulting from the interaction between threats, vulnerabilities, and assets. Security modelling aims to assess and quantify risks to determine appropriate countermeasures.
- **Assets:** Assets are valuable resources or components that require protection, such as data, infrastructure, intellectual property, or personnel. Understanding the importance of assets helps prioritize security efforts.
- **Attack vectors:** Attack vectors represent the paths or methods through which threats can exploit vulnerabilities to compromise a system's security. Security modelling involves analyzing various attack vectors to devise appropriate defences.
- **Controls:** Controls are security measures implemented to reduce the likelihood or impact of threats. They can include technical, administrative, or physical safeguards, such as firewalls, access controls, encryption, policies, and training.
- **Mitigation strategies:** Mitigation strategies involve the actions and countermeasures implemented to reduce risks or minimize the impact of security incidents. Security modelling helps identify and evaluate the effectiveness of different mitigation strategies.
- **Security metrics:** Security metrics are measurable indicators used to assess the effectiveness of security controls, the level of risk, or the impact of security incidents. They provide quantitative or qualitative data for decision-making and improvement efforts.

This output is broadly consistent with recent surveys of security ontologies. Oliveira et al. (2021), for example, report the results of a survey of 57 security ontologies. Their results—presented in Table 3—suggest that vulnerability, asset, threat, countermeasure, risk, and attack are the most commonly

encountered security concepts. A similar set of results was reported by Jarwar et al. (2022; 2022), who directed their attention to the more specific realm of IoT security ontologies. On the basis of this survey work, Jarwar and colleagues proposed an ontology consisting of the following concepts:

- Threat
- Vulnerability
- Security Mechanism
- Asset
- Loss Scenario
- Capability
- Criticality

Table 3. Frequency of concepts in a review of 57 security ontologies (adapted from Oliveira et al. 2021).

Term	Frequency	BFO Mapping
Vulnerability	24	**REALIZABLE ENTITY**
Asset	23	**MATERIAL ENTITY**
Threat	21	
Countermeasure	12	
Risk	9	**PROBABILITY MEASUREMENT INFORMATION CONTENT ENTITY**
Attack	9	**PROCESS**
Attacker	7	**AGENT**
Control	7	
Stakeholder	6	**AGENT**
Consequence	6	**EFFECT**

We will have more to say about these concepts in subsequent sections. For present purposes, however, it is worth noting that some of these concepts present little in the way of a problem as regards their position within a BFO-conformant ontology. Vulnerabilities and capabilities, for example, already feature as part of CCO, where they are both represented as types of **REALIZABLE ENTITY**. Some initial mappings for other concepts are given in Table 3. Blank cells indicate concepts that require further analysis (see below).

4.2 Assets

From a security perspective, an asset is something that is deemed to be worthy of securing. In short, assets are the things we seek to protect because they are, in some sense, valuable to us. What it means for something to be valuable is not particularly clear; nevertheless, there are reasons to think that the notion of value is central to our understanding of risk (see Sales et al. 2018). Boholm and Coverllec (2011, p. 177), for example, suggest that "for an object to be considered 'at risk', it must be ascribed some kind of value." Similarly, the sociologist, Eugene Rosa (1998, p. 28), defines risk as "a

situation or event where something of human value (including humans themselves) has been put at stake and where the outcome is uncertain."

As a means of providing an ontological characterization of the asset concept, we will outline a theory of value that draws inspiration from the so-called active inference approach to brain function (Parr, Pezzulo, and Friston 2022) (see also Section 7.7). According to this approach, we can understand many aspects of brain function as the attempt to minimize the error associated with the prediction of future sensory events. The details of this account need not concern us here; what matters is simply the idea that the brain is engaged in an effort to formulate predictions of the future (specifically, future patterns of neural activity) and then minimize the error associated with those predictions. Crucially, such errors can be minimized in two basic ways. Firstly, the brain can learn about the statistical structure of sensory inputs, so as to improve its capacity to generate (more accurate) predictions. This is the basis of perception (and learning). The second way of minimizing prediction error is to rely on bodily action. The idea here is that one can bring about a predicted state-of-affairs by using one's own actions as an 'operator' that transforms the sensory present into the predicted (sensory) future. To help us get to grips with this idea, consider a situation in which your brain predicts the sensory consequences of reaching for a steaming mug of coffee on your desk. Here, the sensory consequences consist of the particular pattern of proprioceptive input that *would* be obtained if you *were* to implement the relevant reaching movement. Given that you are not, at this very moment, reaching for the mug of coffee, the level of prediction error will be high. You can, however, reduce this error by performing the reaching movement. When the reaching movement is performed, the sensory input will change to match the pattern of predicted sensory input, thereby cancelling out the prediction error. This, in short form, is the active inference approach to action. All action, it is suggested, can be understood as the attempt to minimize prediction error, not by updating one's predictions as is the case with perception, but rather by changing the world to match one's predictions.

A key virtue of active inference is that it provides us with a neurally-plausible approach to understanding a key feature of intelligent systems, namely, the performance of goal-directed action. Consider that, from the standpoint of active inference, a goal is nothing more than a prediction that is poised to entrain the performance of actions. If my goal is to eat, then I can formulate a prediction about a possible future in which I am eating. That prediction is then able to entrain a sequence of actions that culminate in me doing the very thing that fulfils the prediction. My goal, in this case, is nothing more than a particular 'vision' of what I expect the future to be, and my actions are a means of bringing about that future. That is to say, my actions are a means of transforming the actual present into a possible future, specifically, a future that I optimistically expect myself to be in.

This approach to understanding goals serves as the basis for an account of value. Valuable things, we suggest, are those things that play a productive role in enabling us to achieve our goals, which is to say that valuable things are the entities that help us fulfil our predictions regarding the future sensory states that we expect ourselves to be in. Assets are then a proper subset of valuable things. Specifically, we will cast **ASSETS** as **MATERIAL ENTITIES** that are perceived to be valuable by a particular agent (or agent community). Valuable things are the sorts of things that we care about. They are things that are worthy of being shielded from harm. Consider, for example, a simple scenario where I drive my car to the supermarket. My car is an asset because it is a **MATERIAL ENTITY** that enables me to achieve my goal of driving to the supermarket. Because my car is valuable to me, I will implement actions that mitigate the possibility of the car being stolen. I will thus take steps to ensure

the car is locked when it is unattended. If I cared nothing at all for the car, then I would not implement these protective actions.

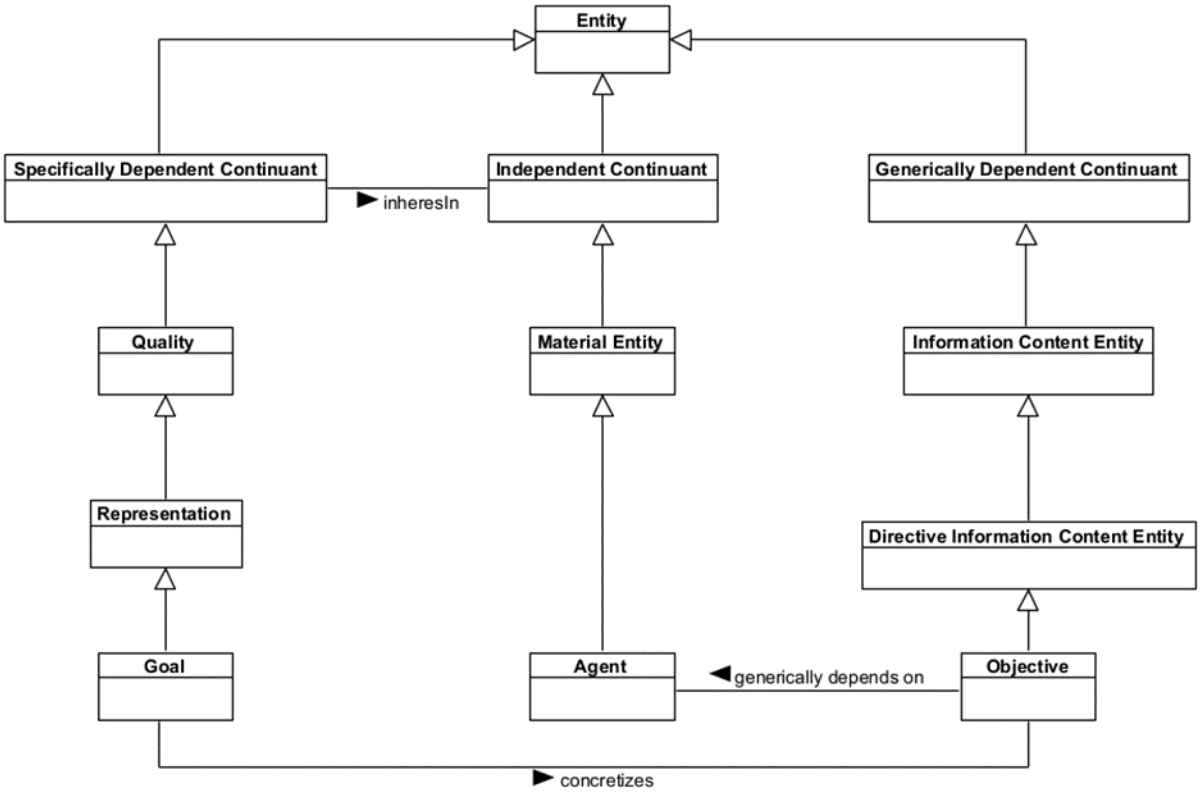

Figure 12. Goals and objectives.

Insofar as we are to understand assets as valuable things, then an ontological characterization of assets will need to include support for the representation of value. What is more, if we are to accept the aforementioned active inference approach to value, then it seems that we will need to include support for a number of other terms, most notably goals, objectives, or expectations about future states. Let us first consider the notion of a goal. What is a goal exactly? According to proponents of active inference, goals can be understood as predictions or expectations about future states-of-affairs (see Clark 2020). From a neuroscientific perspective, such predictions are presumably encoded in a pattern of a neural activity that is to be found in a particular brain region. In this case, the patterns of neural activity represent the thing that is being predicted. That is to say, the patterns of neural activity represent the future state-of-affairs that will be brought about via the implementation of some sequence of actions. What we have here, then, is a basic distinction between the thing-that-is-represented (i.e., the future state-of-affairs) and the thing-that-does-the-representing (i.e., the pattern of neural activity). This distinction is typically understood in terms of the content of a representation (the thing-that-is-represented) and the representational vehicle (the thing-that-does-the-representing). As we saw in Section 3.6, however, however, the appeal to information content typically entails a commitment to a trifold distinction between **GENERICALLY DEPENDENT CONTINUANTS**, **SPECIFICALLY DEPENDENT CONTINUANTS**, and **INDEPENDENT CONTINUANTS**. In particular, the content of a representation is typically understood to be an **INFORMATION CONTENT ENTITY**; the representation, itself, is understood to be a **QUALITY** of

some sort; and these qualities are deemed to ***inhere in*** some **INDEPENDENT CONTINUANT**, such as an **INFORMATION BEARING ARTIFACT**. Using this trifold distinction, we will model **GOALS** as a subclass of **REPRESENTATIONS**. These goals ***inhere in*** agents (or some part thereof), and they are connected to **OBJECTIVES**,[18] which are a type of **INFORMATION CONTENT ENTITY** (see Figure 12).[19]

Goals, we suggest, represent target states-of-affairs. In BFO, these target states-of-affairs are denoted by objectives. For the purposes of the SOfIoTS project, we will introduce a specialization of the **OBJECTIVE** class to refer to situations that an agent aims to bring about courtesy of their own actions. This class is the **SITUATION OBJECTIVE** class. Given that **SITUATION OBJECTIVES** descend from **DIRECTIVE INFORMATION CONTENT ENTITIES** (see Figure 12), we will make use of the ***prescribes*** relation to specify the particular kind of **SITUATION** that a **SITUATION OBJECTIVE** refers to. This raises a question about the nature of situations. Where, exactly, do situations fit within the BFO hierarchy. Our suggestion is that **SITUATIONS** are a type of **OBJECT AGGREGATE**, which is, in turn, a type of **MATERIAL ENTITY**. The constituents of a **SITUATION** are **INDEPENDENT CONTINUANTS** that are connected to a situation via the ***member of*** relation.

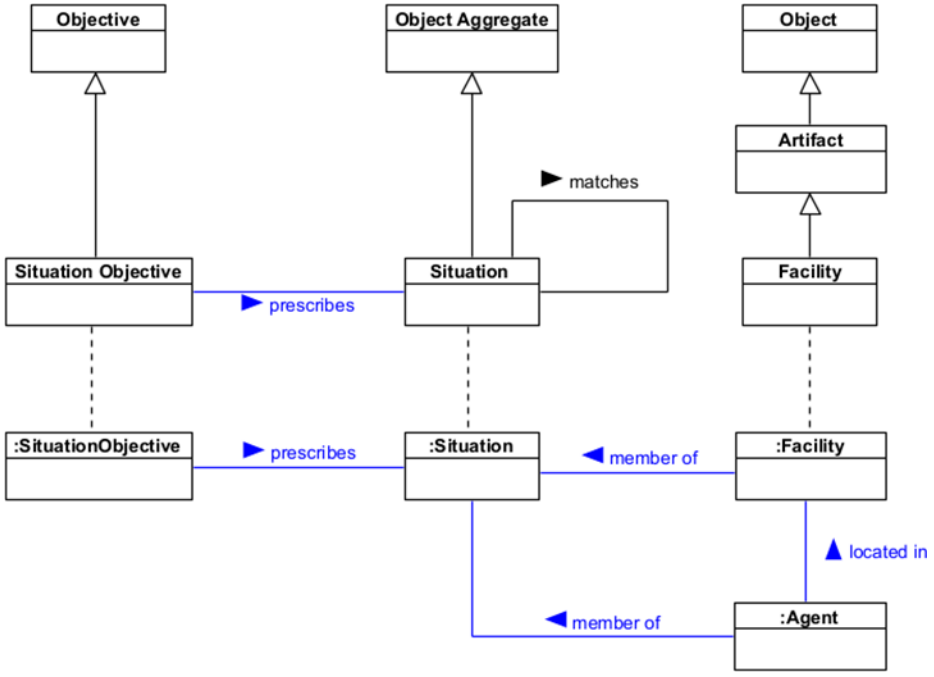

Figure 13. Situation objectives.

Figure 13 shows how instances of the **SITUATION** and **SITUATION OBJECTIVE** can be used to represent target situations. Note that the blue font in Figure 13 signifies that the corresponding

[18] To our mind the use of the term "objective" is a little confusing, since other kinds of directive information content entity are glossed as "specifications." For reasons of consistency, we suggest that future work should consider the relabelling of "objective" to "objective specification."

[19] The extent to which representations ought to be understood as qualities as opposed to information bearing artifacts is admittedly unclear; nevertheless, the present approach is compatible with a number of prior BFO-conformant ontologies (Ceusters and Smith 2010; Limbaugh et al. 2019; Limbaugh et al. 2020; Smith and Ceusters 2015).

relation is of the modal variety. That is to say, it is a relation drawn from the modal relation ontology, which forms part of the CCO suite (see Section 3.4). Also note that **SITUATIONS** can match situations via the ***matches*** relation. This captures the idea that goal fulfilment is a process of matching or aligning situations, such that an actual situation matches a target (predicted) situation.

As a means of making this clear, suppose my goal is to visit the local grocery store. As per the foregoing discussion, my goal will correspond to a prediction about where I will be at some point in the future. Specifically, I predict that I will be at the grocery store. At present, however, I am not at the grocery store. This means that the present situation does not match the predicted future situation. The result is prediction error, which can then be reduced by implementing actions that progressively transform the present situation into the future situation. Suppose that I end up walking to the grocery store. When I arrive at the grocery store, my location will correspond to that depicted in the future situation, and thus the present situation will match the future situation. When this happens, my goal is fulfilled.

Valuable things, we suggest, are those things that play a productive role in enabling us to fulfil our goals. In the foregoing scenario, then, the thing that is valuable to me is *me*, for it is my actions (my walking) that transforms the sensory present into a goal-compliant sensory future. In fact, of all the things that are apt to be valuable to me, the thing that is apt to be of greatest value to me is me. The reason for this is that I am a common feature of all those scenarios in which I rely on my own actions to fulfil my own goals. Without me, there are no more goals that I can fulfil, and there is certainly no way of fulfilling these goals if I am no longer around to fulfil them. So, whatever else we might say about the class of valuable things, it seems likely that agents will be one of the members of this class. This is important, for it helps us make sense of the earlier claim (by Eugene Rosa) that humans, themselves, are one of the things that we might recognize as valuable.

What about other things—things that do not qualify as human agents? Here, it will help to consider a minor modification to the grocery store scenario. Suppose that instead of walking to the grocery store, I decide to use my car to drive to the grocery store. In this case, my car features as one of the things that enables me to achieve my objective. My car thus plays a productive role in enabling me to fulfil (one of my) goals, and it is precisely for this reason, we suggest, that the car qualifies as something that I regard as valuable. Note that this does not mean that my car is valuable in a *tout court* sense. My car is valuable to me, but it may be of little value to you. Indeed, this is highly likely to be the case, for you are not using my car to fulfil your own goals. My car would only be valuable to you if there was some possibility that you could use my car to help you bring about the sorts of things that you expect yourself to bring about. If there is no reason to think that this is the case, then you will not ascribe any sort of value to my car.

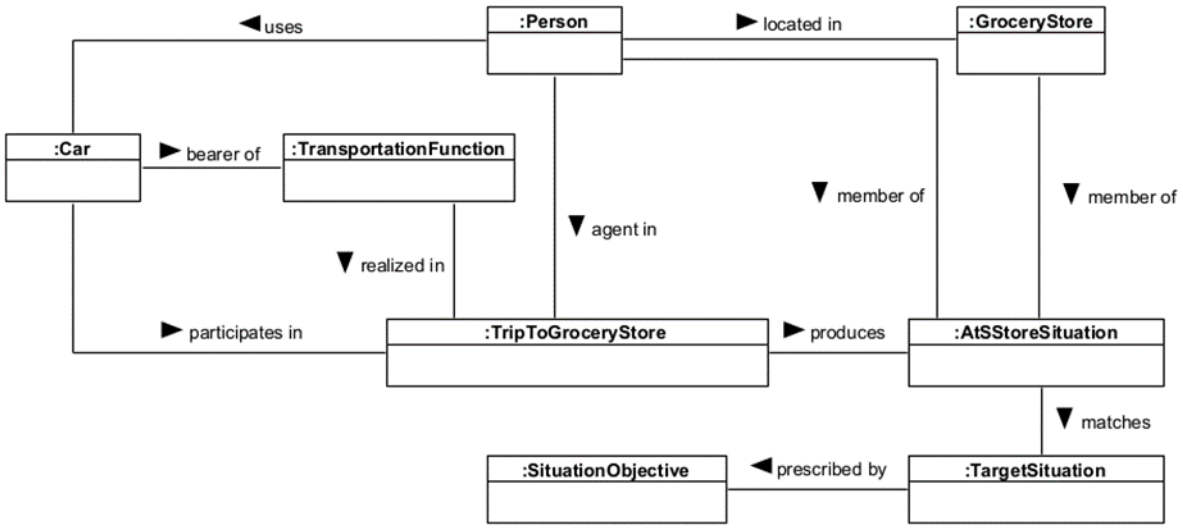

Figure 14. Using resources to perform processes that fulfil objectives.

Figure 14 depicts a scenario in which a person uses a car to drive to the grocery store. The `:TRIPTOGROCERYSTORE` object is an instance of a **PROCESS** that involves both a person and a car as participants. [Note that ***agent in*** is a subtype of the ***participates in*** relation.] Our proposal is that both the person (`:PERSON`) and the car (`:CAR`) count as valuable things on the grounds that they are both participants in a process that culminates in the production of a situation that matches a target situation (i.e., a situation corresponding to a state of goal fulfilment).

We are now in a position to offer some formal definitions of value-related terms:

> **VALUED PROCESS** =*def.* A valued process is a process that produces a situation (or other output) that satisfies (or matches) an objective.
>
> **VALUED SPECIFICALLY DEPENDENT CONTINUANT** =*def.* A valued specifically dependent continuant is a specifically dependent continuant that is involved (as a participant) in a valued process.[20]
>
> **VALUED QUALITY** =*def.* A valued quality is a valued specifically dependent continuant that also qualifies as a quality.
>
> **VALUED REALIZABLE ENTITY** =*def.* A valued realizable entity is a valued specifically dependent continuant that also qualifies as a realizable entity.
>
> **VALUED OBJECT** =*def.* A valued object is an object that either 1) participates in a valued process or 2) is the bearer of a valued specifically dependent continuant (i.e., a valued quality or valued realizable entity).

What is crucial here is the idea that a valued process is a process that produces something that matches an objective, and that this objective is concretized by something that qualifies as a goal. Given this approach to valued processes, we can then define valued objects as things that participate

[20] We will assume that **REALIZABLE ENTITIES** are participants in the processes that realize them.

in these (valued) processes. The upshot is both the person and the car in Figure 14 will qualify as valued objects. What is more, the car's transportation function will count as a valued realizable entity, since it is a specifically dependent continuant that is realized in a valued process.

Now that we have a better understanding of what it means for something to be a valued object, we can turn our attention back towards the notion of an asset. We could, of course, simply state that assets are valued objects. This seems plausible, since valued objects are likely to be the sorts of things that we care about, and thus the sorts of things that are worth securing. On the other hand, we might define an asset as a particular sort of valued object, one that is used by one or more agents for the purposes of completing a task. We could, for example, define an asset as:

> **ASSET** $=_{def.}$ A valued object (o) is an object that is used by an agent (s), where s is an agent in a process (p) that qualifies as a valued process.

This sort of definitional strategy would limit the notion of an asset to the things that are ***used by*** agents in the context of processes that qualify as valued processes. While this strategy is plausible, we remain largely neutral as to whether assets ought to be seen as valued objects or as particular types of valued objects.

The upshot of all this is that assets belong to the metaphysical category of **INDEPENDENT CONTINUANTS** and, more specifically, the category of **MATERIAL ENTITIES**. Assets are material entities that are perceived to be valuable by one or more agents, where the notion of value hinges on the role that a material entity plays in the production of a state-of-affairs that matches an agent's goal.

4.3 Impact, Harm, and Loss

A number of security ontologies include terms pertaining to impact, harm, and loss. Loss events and loss situations, for example, form part of both the Reference Ontology for Security Engineering (ROSE) and the Common Ontology of Value and Risk (COVER) (Oliveira et al. 2022; Sales et al. 2018). In addition, Alanen et al. (2022) have developed an ontology for threat analysis that includes the terms negative impact and loss scenario.

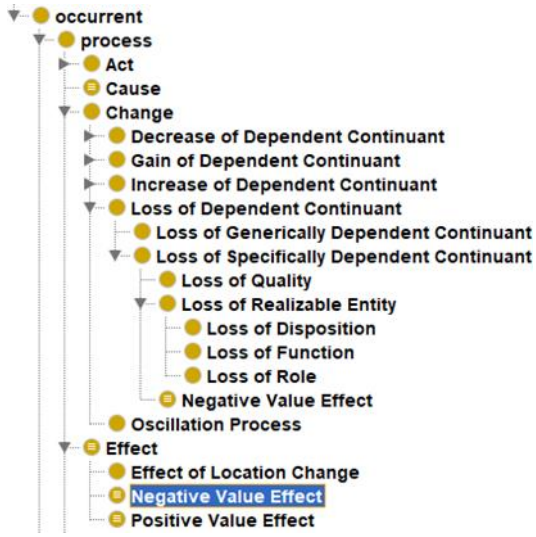

Figure 15. Negative and positive value effects.

Our approach to modelling impact, harm, and loss relies on the participation of valuable things in certain types of processes, or, more generally, occurrents. The CCO includes the **EFFECT** class, which is represented as a type of process (see Figure 15). We introduce two subtypes of the **EFFECT** class, namely, **POSITIVE VALUE EFFECT** and **NEGATIVE VALUE EFFECT**. The definitions of these two effect-related classes are as follows:

> **NEGATIVE VALUE EFFECT** =$_{def.}$ A negative value effect is an effect that involves the loss of a valued property (i.e., a valued specifically dependent continuant).
>
> **POSITIVE VALUE EFFECT** =$_{def.}$ A positive value effect is an effect that involves the gain of a valued property (i.e., a valued specifically dependent continuant).

These definitions draw on the idea that an effect can be a process that entails the gain or loss of some property. In Figure 15, for example, we see that **NEGATIVE VALUE EFFECTS** are a subclass of both the **EFFECT** class and the class **LOSS OF SPECIFICALLY DEPENDENT CONTINUANT**, which is, in turn, represented as a type of **CHANGE**. In CCO, changes are defined as follows:

> **CHANGE** =$_{def.}$ A process in which some independent continuant endures and 1) one or more of the dependent entities it bears increase or decrease in intensity, 2) the entity begins to bear some dependent entity or 3) the entity ceases to bear some dependent entity.

Figure 16 illustrates how change processes can be used to represent the gain and loss of a specifically dependent continuant, in this case, a role.[21] Here we see that roles are gained via their participation in a particular kind of change process, namely, **GAIN OF ROLE**. Likewise, roles are lost via their participation in another sort of change process, namely, **LOSS OF ROLE**. A **NEGATIVE VALUE EFFECT**, recall, is a type of process (specifically, a type of **EFFECT**) that is also a type of loss-related change process. In particular, it is both an **EFFECT** and a **LOSS OF SPECIFICALLY DEPENDENT**

[21] This example is adapted from CUBRC (2020).

CONTINUANT. The things that participate in **NEGATIVE VALUE EFFECTS** are thus **SPECIFICALLY DEPENDENT CONTINUANTS**, which includes things such as roles, capabilities, qualities, dispositions, abilities, tendencies, and so on. When these things ***participate in*** loss-related change processes, they no longer ***inhere in*** the **INDEPENDENT CONTINUANTS** that bear them. This is important when it comes to what we earlier referred to as **VALUED SPECIFICALLY DEPENDENT CONTINUANTS**. If a **VALUED SPECIFICALLY DEPENDENT CONTINUANT** ***participates in*** a loss-related change process (i.e., a type of **LOSS OF DEPENDENT CONTINUANT** process), then the relevant process is one in which something of value is being lost. This is the basis of our understanding of what a **NEGATIVE VALUE EFFECT** is. In essence, a **NEGATIVE VALUE EFFECT** is a process in which some **VALUED SPECIFICALLY DEPENDENT CONTINUANT** is lost as a result of its participation in the relevant process. Conversely, a **POSITIVE VALUE EFFECT** is a process in which some **VALUED SPECIFICALLY DEPENDENT CONTINUANT** is gained as a result of its participation in the relevant process.

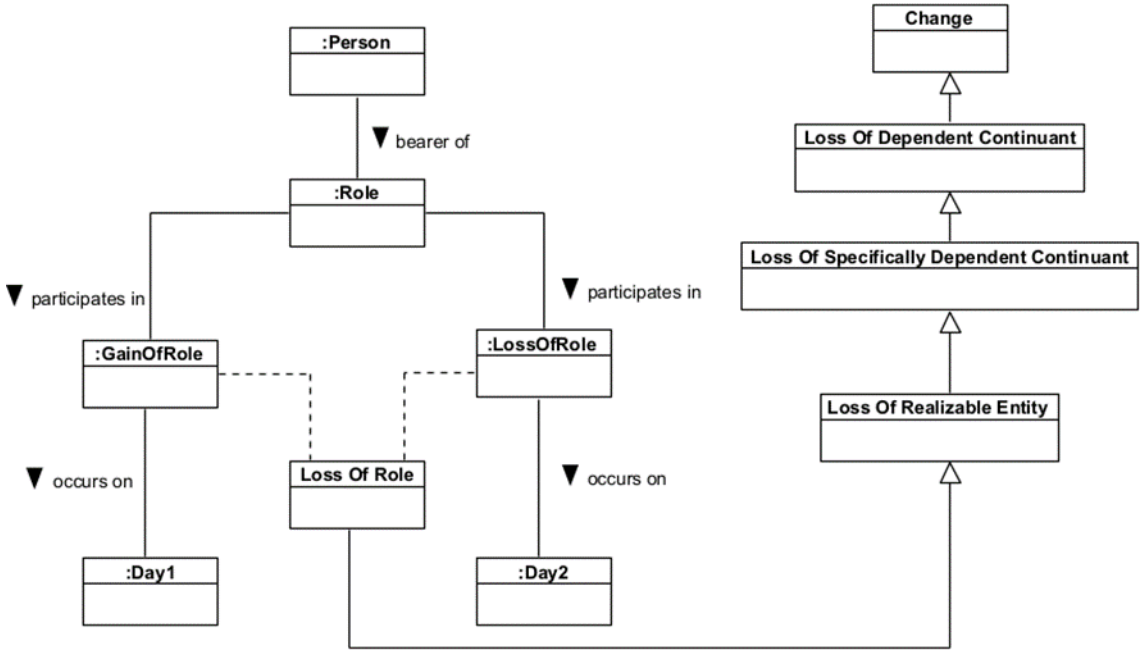

Figure 16. Representing information about the gain and loss of an occupational role. [Dashed lines represent instantiation (instance-of) relations.]

Note that we are talking here of loss-related change processes and gain-related change processes. The CCO, however, features other types of change processes. These include processes that reflect the increase or decrease of a dependent continuant (see Figure 15). While we will not discuss these processes, it seems likely that decreases (or increases) in some **VALUED SPECIFICALLY DEPENDENT CONTINUANT** are also relevant to our understanding of what makes something a **NEGATIVE VALUE EFFECT** (or **POSITIVE VALUE EFFECT**).

NEGATIVE VALUE EFFECTS and **POSITIVE VALUE EFFECTS** are both examples of what we might call "impactful effects." They are impactful in the sense that they have a bearing on things that matter to us. If I value my car on the grounds that it has a certain capability or functionality (e.g., a

transportation function), then the loss of that capability/function is something that matters to me. Similarly, I might value my smartphone on the grounds that it enables me to access online information or communicate with others. In this case, the value of the smartphone is tied to the features of the smartphone and the way those features relate to my own goals (or, in the language of active inference, my high-precision, optimistic expectations). [Contemporary smartphones are, of course, multi-functional devices, in the sense that they enable us to do many things. In addition to accessing online information, they also allow us to take photos, record videos, and participate in a variety of communication-related activities.]

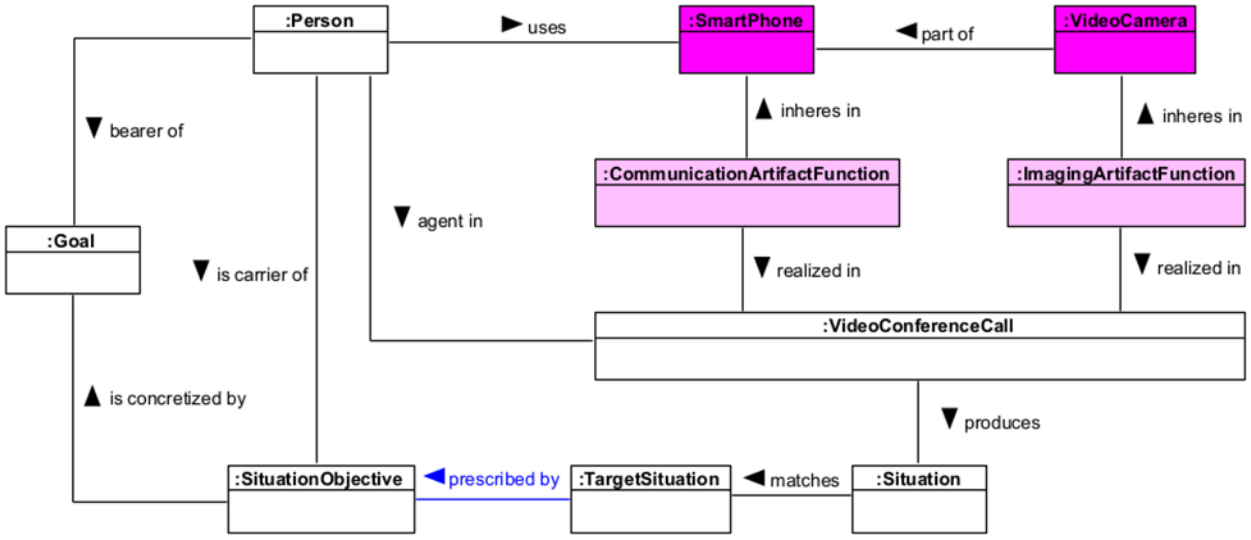

Figure 17. The valorization of specifically dependent continuants (light pink shading) and the valued objects (dark pink shading) they inhere in.

Figure 17 illustrates one way that the various functions of a smartphone device might acquire value to a user. In this case, the two artifact functions (`:COMMUNICATIONARTIFACTFUNCTION` and `:IMAGINGARTIFACTFUNCTION`) are deemed to be valuable to `:PERSON` on the grounds that they enable `:PERSON` to bring about (or produce) states-of-affairs (or **`SITUATIONS`**) that match some target situation, thereby fulfilling a **`SITUATION OBJECTIVE`**.

In BFO, functions are typically understood to be dispositions. Dispositions are, in turn, understood to be internally-grounded properties (or features) of an entity that are realized in occurrents (Arp and Smith 2008; Goldfain, Smith, and Cowell 2010; Hastings et al. 2011). In philosophy, a canonical example of a disposition is fragility (Mumford 1998). Thus, we might say that a vase is fragile on the grounds that it is disposed to break when it is dropped onto a hard surface. In BFO, we would say that fragility is a particular sort of disposition that inheres in objects of a certain sort; in this case, a vase. Smartphones are perhaps somewhat more robust than vases, but they are nevertheless disposed to break in certain situations. If, for example, we accidentally drop a smartphone while leaning over a balcony, then the smartphone (or some its components) may be disposed to break. If this sort of process unfolds, then we will end up in a situation where certain properties of the smartphone are no longer borne by the smartphone. Prior to the dropping process, the smartphone (and its constituent camera) may have possessed functionalities that are valuable to us (see Figure 17). Once the dropping process has concluded, however, these functionalities are apt to be lost. The loss of these functionalities has an impact that goes beyond the mere physical forces associated with

the phone–ground interaction. In particular, the loss of what we earlier dubbed **VALUED SPECIFICALLY DEPENDENT CONTINUANTS** means that certain routes into the future are now blocked. Prior to the smartphone breaking, we were able to use the smartphone to bring about (or produce) certain future states that matched the states that we (optimistically) expect ourselves to be in (or to bring about courtesy of our own actions). Once the smartphone is broken, however, we are no longer able to bring about (or generate) these futures. In essence, our future is constrained in a way that it was not constrained before. This idea of the future being affected by the loss of valued specifically dependent continuants is, we suggest, key to our understanding of a variety of impact-related concepts.

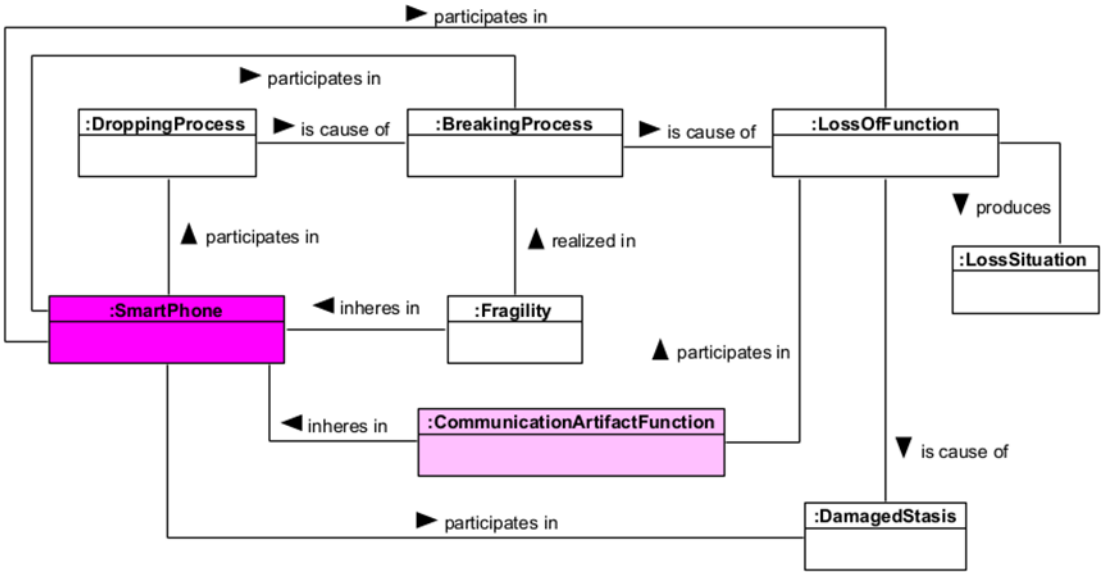

Figure 18. Dropping a smartphone causes it to break, which entails the loss of one or more valued specifically dependent continuants.

Let us consider how the aforementioned smartphone dropping case might be represented in CCO. The key features of this case are depicted in Figure 18. There are actually four processes at work here: `:DROPPINGPROCESS`, `:BREAKINGPROCESS`, `:LOSSOFFUNCTION`, and `:DAMAGEDSTASIS`. These processes form part of a causal chain that (in this case) begins with the `:DROPPINGPROCESS` and concludes with the `:DAMAGEDSTASIS` process. The `:SMARTPHONE` participates in all these processes. The phone is the bearer of a `:FRAGILITY` property, which is an instance of the BFO **DISPOSITION** class. The `:FRAGILITY` of the `:SMARTPHONE` is ***realized in*** the `:BREAKINGPROCESS`, which is when the `:SMARTPHONE` (or some part thereof) breaks. This is followed by the `:LOSSOFFUNCTION` process, which has `:COMMUNICATIONARTIFACTFUNCTION` as one of its participants. As we saw in Figure 17, `:COMMUNICATIONARTIFACTFUNCTION` is an instance of **VALUED SPECIFICALLY DEPENDENT CONTINUANT**, so the loss of this function has a value-related impact. The upshot is that a machine reasoner will recognize that `:LOSSOFFUNCTION` is a **NEGATIVE VALUE EFFECT** on the grounds that it is an **EFFECT** that entails the loss of something that has previously been classified as a **VALUED SPECIFICALLY DEPENDENT CONTINUANT**. From Figure 18, we can see that the `:LOSSOFFUNCTION` process produces a situation named `:LOSSSITUATION`. The naming of this instance (or individual) is, of course, of little consequence from an inferential perspective. Nevertheless, we can now define a class that represents

states-of-affairs in which something of value has been lost. Call this a **LOSS SITUATION**. **LOSS SITUATIONS** are defined as follows:

> **LOSS SITUATION** =*def.* A loss situation is any situation that is produced by a negative value effect.

At the outset of this section, we noted that a number of ontologies feature terms like loss event, loss situation, and loss scenario (Alanen et al. 2022; Oliveira et al. 2022). The **LOSS SITUATION** class provides us with means of understanding these terms from a BFO/CCO perspective.

LOSS SITUATIONS are, of course, inherently negative, but there is no reason why we cannot include a positive counterpart to such situations. Such situations are what we will call **GAIN SITUATIONS**. These are defined as follows:

> **GAIN SITUATION** =*def.* A gain situation is any situation that is produced by a positive value effect.

LOSS SITUATIONS and **GAIN SITUATIONS** are both situations that have some sort of impact on a particular agent. In short, such situations are what we might call **IMPACTFUL SITUATIONS**. Such situations can be defined as the union of **LOSS SITUATION** and **GAIN SITUATION**. **IMPACTFUL SITUATIONS** will then be those situations that are produced either by **NEGATIVE VALUE EFFECTS** or **POSITIVE VALUE EFFECTS**. These effects are what we might call negative and positive impacts, although we have made no attempt to represent the notion of an impact event or process in CCO. To our mind, impacts are simply 'effects that matter', and what that means, at least relative to the foregoing analysis, is that negative (or positive) impacts are simply another term for negative (or positive) value effects.

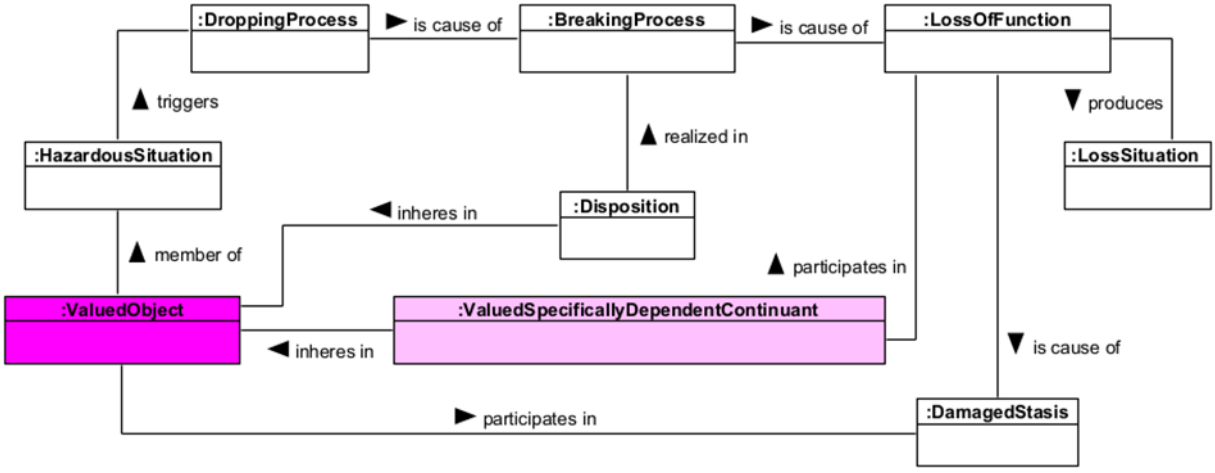

Figure 19. Hazardous situations.

Thus far, we have defined three types of situations, namely, **LOSS SITUATIONS**, **GAIN SITUATIONS**, and **IMPACTFUL SITUATIONS**. Quite plausibly, however, an ontology of risk and security ought to consider situations of other types. Some of these situations will emerge in subsequent sections; for present purposes, however, let us consider how the notion of a hazardous situation might be accommodated by CCO. One approach is depicted in Figure 19. Figure 19 shows

an abstract version of the smartphone dropping case, which includes an individual named :HAZARDOUSSITUATION. This individual is an instance of the class **HAZARDOUS SITUATION**, which is defined as follows:

> **HAZARDOUS SITUATION** =*def.* A hazardous situation is any situation that triggers a negative value effect (either directly or indirectly).

This definition features a new term, namely, the ***triggers*** relation. This term is derived from work on both the COVER and ROSE ontologies (Sales et al. 2018; Oliveira et al. 2022). In essence, a situation is said to trigger a process, when that situation features conditions that trigger the exercise or manifestation (or, in BFO terms, the realization) of realizable entities.[22] In short, the idea is that some situations are conducive to the instantiation of certain types of processes. In the smartphone dropping case, this situation could be one in which the smartphone owner is handling the smartphone whilst leaning over a balcony. This is a situation in which the :DROPPINGPROCESS, if it should occur, will culminate in a :LOSSSITUATION, which, as we have already discussed, is a **NEGATIVE VALUE EFFECT**. Given the aforementioned definition of a hazardous situation, this will result in a machine reasoner inferring that :HAZARDOUSSITUATION must be an instance of **HAZARDOUS SITUATION**. For more on trigger relations and the notion of a disposition trigger, see Section 6 (see also Ray et al. 2016).

4.4 Risk

While the notion of risk is central to security modelling, its ontological characterization remains obscure. As noted above, a popular definition of risk is owed to the sociologist Eugene Rosa (1998). Rosa suggests that risk is "a situation or event where something of human value (including humans themselves) has been put at stake and where the outcome is uncertain." This seems to suggest that risk is, at root, a situation or event. That, however, doesn't seem quite right. While we might talk of a risky situation and/or risky process, it is not clear that risk itself ought to be understood as either a situation or a process. Earlier we suggested that a situation is an object aggregate or material entity, which puts situations in the metaphysical category of independent continuants. Processes (or events), by contrast, belong to the metaphysical category of occurrents. The problem is that risk seems to be neither an independent continuant nor an occurrent. Risk is not a material entity, like a chair, a person, or a cat; but nor it is something that unfolds through time, like an occurrent. If there is a risk of X occurring, then the actual occurrence of X does not radically change the risk. To be sure, once X occurs, then the risk is either realized or unrealized, but it seems perfectly plausible that the risk associated with X could exist prior to the actual occurrence of X. If that is true, however, then risk cannot be a process, for processes (as occurrents) cannot exist prior to their actual occurrence.

Given this, it seems unlikely that risk ought to be understood as either an occurrent or an independent continuant. This means that risk must be a dependent continuant—either a specifically or generically dependent continuant—for that is the only category of entities that remain once occurrents and independent continuants have been eliminated.

Some insight into the ontic character of risk is provided by a survey of earlier work. Of particular interest is the analysis offered by Oliveira et al. (2022). They begin with the aforementioned

[22] From a philosophical standpoint, such conditions might be referred to as trigger conditions (see McKitrick 2018).

definition by Rosa and identify three necessary and sufficient conditions that are entailed by this definition. These are as follows:

1. **Interest Condition:** Risk relates to some possible state of reality that affects someone's interest, either positively or negatively.

2. **Uncertainty Condition:** Risk involves uncertainty about whether or not such a state will hold in the future; thus, if an event is certain to happen (such as the sun rising tomorrow), one cannot ascribe a risk to it.

3. **Possibility Condition:** Risk is about a possible state of reality (thus ruling out the possibility of talking about the risk of someone turning into a werewolf).

These conditions suggest that risk refers to a possible, albeit uncertain, future state-of-affairs that has some sort of value-laden impact (either positive or negative) for at least one agent or agent collective. The notion of uncertainty looks to be particularly important to our understanding of risk, for if a future state-of-affairs is guaranteed to occur, then there seems little reason to resort to risk-related terminology. If, as Oliveira et al. (2022) state, we already know that the sun will rise tomorrow, then there is no risk associated with this particular occurrence (or non-occurrence). This suggests that risk is tied to issues of positive epistemic standing. If one *knows* that X, then there is no risk associated with X. Accordingly, the more one knows about X, the less risk there is associated with X. In short, knowledge (or the acquisition of knowledge) is a means of reducing risk. This is not to say that all forms of knowledge are conducive to risk reduction, however. Consider that one could know the probabilities associated with the occurrence of two mutually exclusive events, but this knowledge need not diminish risk. This point is made by Parr et al. (2022):

> Risk, a common notion in economics, corresponds to the fact that there can be a one-to-many mapping between policies [action strategies] and their consequences—in the sense that one can obtain several different outcomes (by chance) under the same policy. One example is a gambling scenario with stochastic rewards (e.g., a one-armed bandit, aka a slot machine), wherein one could know the reward distribution—say, that one will obtain reward 10 percent of the time. This is called a risky situation in economics because, after the same move (pulling a lever), one could obtain two different observations (reward or no reward). This means one has to choose policies or plans that accommodate uncertainty. (Parr, Pezzulo, and Friston 2022, p. 35)

Here, it seems that one could possess knowledge about the likelihood of a certain outcome (e.g., there is a 10% chance of a reward), but this knowledge need not eliminate the risk of participating in a process (e.g., inserting a coin and pulling the lever).

There are a couple of things we can glean from this. The first is that risk is referring to things that may or may not happen in the future. That suggests that our ontological approach to risk will need to draw on the modalistic components of CCO. In particular, we will need to utilize relations that form part of the modal relation ontology (Jensen et al. 2018) (see also Section 3.4). A second insight is that risk belongs to the realm of 'cognitive' constructs, such as beliefs, knowledge, interpretations, appraisals, estimates, judgements and the like. In particular, it seems plausible that risk is something akin to a belief regarding the likelihood or probability of certain things occurring. Likelihood is, in fact, a central feature of some ontological approaches to modelling risk. The ROSE ontology, for example, refers to two kinds of likelihood: triggering likelihood and causal likelihood (Oliveira et al.

2022). According to Oliveira et al. (2022) triggering likelihood is the likelihood of an event (or process) occurring within a given situation. Causal likelihood, by contrast, refers to the probability that one event (or process) will lead to another event (or process):

> Triggering Likelihood inheres in a Situation Type, and it refers to how likely a Situation Type will trigger an Event Type once a situation of this type is brought about by an event; the Causal Likelihood inheres in an Event Type, and it means the chances of an event causing, directly or indirectly, another one of a certain type. (Oliveira et al. 2022, p. 369)

As a means of helping us understand how to deal with risk from a BFO perspective, let us begin by representing these two forms of likelihood. Figure 20 demonstrates how the notion of causal likelihood could be represented in a BFO-conformant manner. Here, we have introduced a new class, called **CAUSAL LIKELIHOOD**. This is represented as a type of **PROCESS LIKELIHOOD**, which is, in turn, a type of **PROBABILITY MEASUREMENT INFORMATION CONTENT ENTITY**. Instances of this class are intended to represent the likelihood or a process occurring. Specifically, the class is defined as follows:

> **PROBABILITY MEASUREMENT INFORMATION CONTENT ENTITY** $=_{def.}$ A Measurement Information Content Entity that is a measurement of the likelihood that a Process or Process Aggregate occurs.

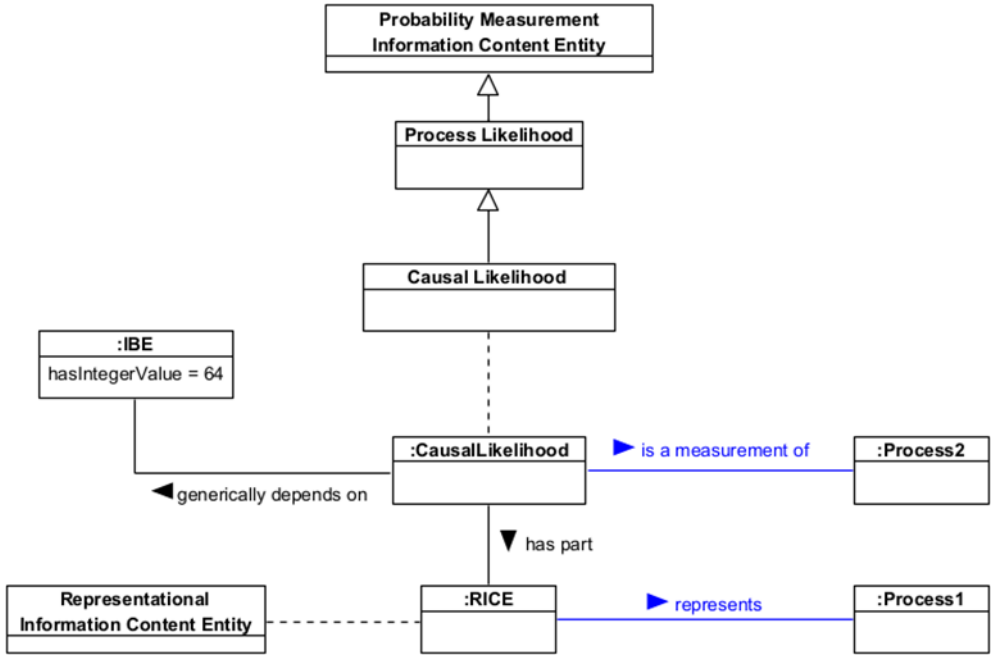

Figure 20. A causal likelihood is a process likelihood that specifies the likelihood that one process will follow another. The causal likelihood is the likelihood (or probability) of a causal relationship between two processes (:PROCESS1 and :PROCESS2), where :PROCESS1 is the direct or indirect cause of :PROCESS2. The likelihood is represented by a representational information content entity (:RICE) that forms part of a causal likelihood information content entity.

In Figure 20, we see that the likelihood information is represented by `:CAUSALLIKELIHOOD`. The likelihood value is represented by the value associated with `:IBE`, which is an **INFORMATION BEARING ENTITY**. The process that is being referred to by the likelihood information is

:PROCESS2, which is linked to :CAUSALLIKELIHOOD via the ***is a measurement of*** relation. Note that this relation appears in blue font, which means it is drawn from the modal relation ontology. :PROCESS2 is thus a *bona fide* instance of the **PROCESS** class, but it does not represent (or refer to) any particular process. That is to say, there is no actual process in reality that we can identify as being the referent of :PROCESS2. What the ***is a measurement of*** relation means, in this context, is that processes that are identical to :PROCESS2, or at least very similar to :PROCESS2, will have a certain likelihood of occurring. In Figure 20, this likelihood is 64%, which is represented as an integer value. There is, however, no reason why the likelihood value is restricted to the realm of integers; an alternative scheme could rely on the use of categorical values to represent likelihood (e.g., low, medium, and high).

While Figure 20 focuses on causal likelihoods, the ***is a measurement of*** relation is, in fact, common to both causal likelihoods and triggering likelihoods. What distinguishes causal likelihoods from triggering likelihoods is the entity that is represented by one of the components (or parts) of a **PROCESS LIKELIHOOD**. In Figure 20, this component is labelled :RICE, which is a type of **REPRESENTATIONAL INFORMATION CONTENT ENTITY**. The thing that is represented by :RICE is, in this case, :PROCESS1, which is another process. The idea here is that :PROCESS1 is causally connected to :PROCESS2, such that :PROCESS1 will cause :PROCESS2 to occur. This provides us with a basic means of representing causal likelihoods.

What about triggering likelihoods. Our proposal is that triggering likelihoods can be modelled in precisely the same way as causal likelihoods. The only difference here relates to the type of entity that is targeted by the ***represents*** relation. In the case of causal likelihoods, this is a processual (or, more generally, an occurrent) entity. For triggering likelihoods, the relevant entity will be an instance of the **SITUATION** class. This enables us to represent the likelihood that a particular process will occur in a given *situation*, as opposed to the idea that a given process will be caused by a given *process*.

As with the ***is a measurement of*** relation, the ***represents*** relation is depicted in blue font in Figure 20. This indicates that it is a modal variant of the standard represents relation. Accordingly, :PROCESS1 is not a real process; it is instead, a placeholder for actual processes that are identical to (or at least highly similar to) :PROCESS1. [The same is true of situations in the case of triggering likelihoods.] One way of thinking about these processes is as process *types*. That is to say, :PROCESS1 and :PROCESS2 refer to general categories of processes, in the same way that a conventional (programmatic) class refers to a category of entities. This may seem a little counterintuitive, for :PROCESS1 and :PROCESS2 are clearly instances of classes: they are instances of a type, not a type themselves. Despite this, it is typically assumed that risk models benefit from the ability to treat instances as types. Oliveira et al. (2022), for example, talk of situation types and event types as instances that support the representation of likelihood information. This presents something of a problem for BFO, for BFO insists on a strict demarcation between types (or universals) and instances (or particulars) (see Arp et al. 2015). Accordingly, we are not permitted to mix and match these entities, such that types are represented as instances and vice versa.

One way round this problem is to do what we have done here and simply assume that the entity referenced by a modal relation is best regarded as a type or category of things, as opposed to an actual instance of a thing. There are a number of ways this sort of assumption could be made more explicit. One approach is exemplified by the C3O. The C3O includes a class labelled **OPEN SYSTEMS**

INTERCONNECTION CATEGORY, which is represented as a sub-class of **NOMINAL MEASUREMENT INFORMATION CONTENT ENTITY**. The instances of **OPEN SYSTEMS INTERCONNECTION CATEGORY** are then the actual categories that fall under this heading (e.g., application layer, network layer, physical layer, and so on). In a practical data modelling context, these instances could be used to represent the fact that some entity belongs to this category. A somewhat similar scheme is adopted by Hagedorn et al. (2019), as part of their effort to incorporate Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT) categories into an ontology for additive manufacturing. Their approach involves representing SNOMED CT categories as information content entities that are used to denote the classification of entities that fall within that category.

An alternative way of thinking about `:PROCESS1` and `:PROCESS2` is as prototypical instances of the classes from which they are instantiated. This is the sort of approach that is adopted by the OBO relation ontology.[23] In particular, the OBO relation ontology defines a ***has prototype*** relation, which is defined as follows:

> ***has prototype*** $=_{def.}$ x has prototype y if and only if x is an instance of C and y is a prototypical instance of C. For example, every instance of heart, both normal and abnormal is related by the has prototype relation to some instance of a "canonical" heart, which participates in blood circulation.

Further discussion of this issue would take us too far afield; nevertheless, it should be relatively clear that category-like information can be represented in BFO, even if the precise details of the representational strategy require further delineation.

Having explored the approach to representing likelihood information, we are now in a position to define some risk-related terms. A useful starting point is the notion of a risky situation. Intuitively, a risky situation is a situation in which there is a certain likelihood of a particular process occurring. That process is, we suggest, one that has some sort of impact on an agent, specifically, the agent that is exposed to risk. The discussion in Section 4.3 provides us with a means of understanding impact. An **IMPACTFUL SITUATION**, we suggested is either a **LOSS SITUATION** or a **GAIN SITUATION**, with the former resulting from a **NEGATIVE VALUE EFFECT** and the latter resulting from a **POSITIVE VALUE EFFECT**. A risky situation can thus be understood as a situation in which there is a certain likelihood of a value-related effect occurring. For the most part, risky situations will be those in which an individual stands to lose something, which is to say that the relevant form of value-related effect is a negative value effect. The upshot is that risky situations can be defined as follows:

> **RISKY SITUATION** $=_{def.}$ A risky situation is a situation in which there is a certain likelihood of a negative value effect occurring.

This definition is, admittedly, a little vague, for we have not specified what sort of likelihood qualifies as a "certain likelihood." All we can really say here is that beyond some predetermined threshold (e.g., 50%), a situation will be classed as risky. The actual value of this threshold is apt to vary on a case by case basis, and there is, as far as we can tell, no universal likelihood value that distinguishes risky situations from their non-risky counterparts.

[23] The same sort of approach seems to be adopted by Jensen et al. (2018) when they talk of prototypical sensors being involved in prototypical sensor processes.

As a means of exemplifying all this, consider a state-of-affairs in which some valued object (e.g., a laptop computer) is left unattended. Suppose you are in a coffee shop and wish to purchase a refill. The question is whether you ought to leave your laptop unattended while you visit the counter. Here, your choice may very well be informed by the nature of the situation in which you find yourself. Is it generally ok to leave things unattended in a coffee shop environment? Perhaps that depends on which coffee shop you are in, where the coffee shop is, whether you have been recently exposed to a newspaper article reporting on a spate of coffee shop-related thefts, and so on. Depending on the situation, you will no doubt arrive at different estimates as to the likelihood of your laptop being stolen. This likelihood estimate is a form of triggering likelihood. The situation referred to by the ***represents*** relation is a situation that matches the situation in which you currently find yourself. It is perhaps closely related to a prototypical situation (recall the above discussion). The thing that is more or less likely in this situation is your laptop being stolen. Ontologically, this is an occurrent entity, specifically, a process. If this process should occur in your absence, then you will be left in a situation where your laptop is no longer available to you. You will no longer have access to the laptop or the various things that it enables you to do. In this sense, then, you have lost something valuable—you will have lost a valuable thing. The loss of that thing leaves you in a new situation, namely, a loss situation.

In addition to risky situations, it seems plausible that processes are also the sorts of things that could be deemed risky. The extent to which these are genuinely distinct from risky situations is, admittedly, a little unclear, for it could be argued that the transition from one situation to another situation can only occur via the instantiation of some sort of process. Accordingly, all forms of risky situations could be reduced to risky processes. In the coffee shop scenario, for example, there is nothing inherently risky about the fact that you are currently sitting in front of your laptop, contemplating the possibility of ordering another coffee. What makes the situation risky (or non-risky) is what you decide to do in this situation. If you decide to leave your laptop unattended, then you are participating in a process that exposes you to the risk of your laptop being stolen, and it is only once this process is performed that the notion of a risky situation has any traction. Given this, we might want to redirect our attention to the processes that could be performed in a given situation, as opposed to the actual situation in which such processes might occur. As we said, the extent to which we can discriminate between these perspectives is not particularly clear-cut. Nevertheless, it does seem important that we are able to talk about risky processes and not just risky situations.

With this in mind, we can define a risky process as follows:

> **RISKY PROCESS** =*def.* A risky process is a process that has a certain likelihood of causing a negative value effect.

Once again, the appeal to "certain likelihood" is uncomfortably vague, but the general idea of a risky process as one that has a certain likelihood (or probability) of causing a negative value effect ought to be largely uncontroversial. In this case, the relevant likelihood is represented by instances of the **CAUSAL LIKELIHOOD** class. In the coffee shop case, the counterpart to `:PROCESS1` in Figure 20 is the process of visiting the counter and thereby leaving your laptop unattended. The counterpart to `:PROCESS2` is the process of the laptop being stolen. What you decide to do will no doubt depend on how likely you deem `:PROCESS2` to be given the fact that `:PROCESS1` occurs.

The astute reader will no doubt be aware that we are glossing over a number of complexities here. For a start, your decision to leave the laptop unattended will be based on your assessment of the risk

of leaving the laptop unattended, but in evaluating this risk, you are not actually performing the relevant processes. You are not, for example, actually going to the counter and then evaluating the risk. Rather, you are contemplating whether or not to perform this process (in the future). In this sense, then, it cannot be correct to say what is being evaluated is the risk attached to a concrete process. The foregoing discussion suggests that `:PROCESS1` and `:PROCESS2` are process instances, even if they are non-actual processes. This is, at least, plausible when it comes to our representation of likelihood-related information, but the use of this information as part of some decision-making process seems to suggest that the processes that are being compared to the non-actual (or prototypical) instance of a process are not processes either! Again, when you assess the risk of leaving your laptop unattended, you are not actually leaving your laptop unattended, you are merely contemplating the possibility of leaving your laptop unattended. To be sure, your contemplative acts do count as a genuine process, but this process is more of a cognitive process or a thinking process; it is not the actual process about which you are thinking.

All of this suggests that there is much more work to be done in respect of risk, its relationship to processes of various sorts, and the role that risk plays in risk assessment processes. We will not attempt to resolve these issues here, for they are deserving of a more detailed analysis than the one that can be offered here. Despite this, we hope to have provided at least the rudimentary basis for the representation of risk-related terms in a manner that conforms with the architectural principles of the BFO framework.

4.5 Threats and Attacks

Talk of risk often goes hand-in-hand with talk of threats and attacks, at least in security settings. To some extent, this is unsurprising, for the sorts of risks that arise in cybersecurity contexts are typically ones that are connected to the possibility of some sort of cyber-attack, and such attacks are often perpetrated by actors that are plausibly understood to be threats. The extent to which threats are an invariable feature of risky situations remains unclear, for it seems possible that one could be exposed to risk in the absence of a discernible threat. Consider the case of the casino gambler who stakes his life savings on a particular outcome. There is, it seems, some sort of risk here, but it is much less clear that there is anything resembling a threat. On the other hand, we might say that the gambler is a threat to himself, for it is gambler that exposes himself to the possibility of an unwelcome future.

In an effort to keep things simple, let us first consider the notion of an attack. From an ontological perspective, an attack is relatively easy to classify. All attacks are processual entities, and they thus belong to the metaphysical category of occurrents. To be a little more specific, we can state that attacks are deliberate or intentional processes. That is to say, they are processes that are deliberately performed by agents with the express goal of inflicting harm. This is broadly consistent with the way that cyber-attacks are represented in the C3O. In particular, **`CYBER ATTACKS`** are represented as types of **`CYBER ACTS`**, which are, in turn, types of **`ACTS`**. An **`ACT`** is then understood to be a **`PROCESS`** "in which at least one agent plays a causative role." Figure 21 illustrates the types of cyber-attacks represented in the C3O. Cyber-attacks are defined as follows:

> **`CYBER ATTACK`** $=_{def.}$ Cyber Attack is a cyber act that is malicious and directed at a portion of cyberspace.

As things stand, there is much that is incomplete about this definition. For a start, the mere idea of an agent playing a causative role in a process does not seem sufficient to distinguish the realm of

intentional/deliberate acts from the realm of mere non-intentional processes. To make matters worse, the CCO includes a class labelled **UNPLANNED ACT**, which is defined as:

> **UNPLANNED ACT** $=_{def.}$ An Act in which at least one agent plays a causative role and which is not prescribed by some objective held by any of the Agents.

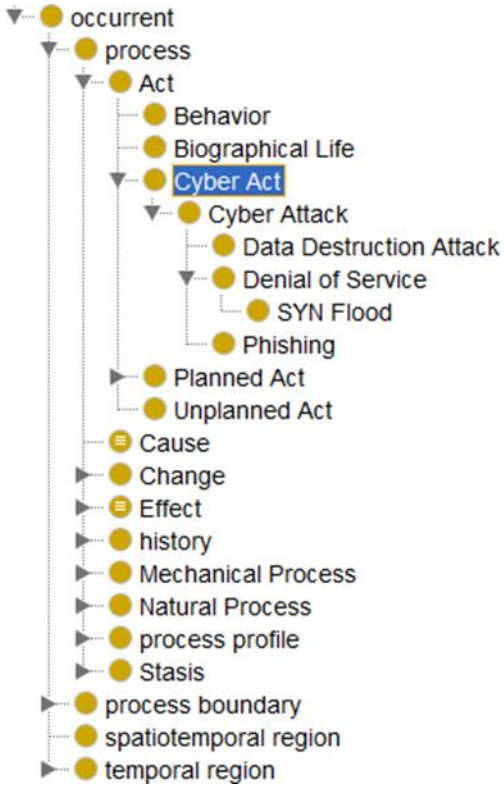

Figure 21. Cyber act taxonomy in the C3O.

Unplanned acts, it seems, are types of acts, but the thing that distinguishes an unplanned act from a planned act is the fact that one or more agents that are causally implicated in the process have some sort of objective. In this sense, then, a cyber-attack seems to be more of a planned act than a mere act. What it means for an agent to play a causative role in a process is, unfortunately, not something that is explicated by either the CCO or the C3O. To the best of our knowledge, the notion of a causative role means that a process was caused to occur by an agent. That, however, opens the door to unintentional or accidental acts (e.g., the accidental dropping of a vase), which are not typically thought of as acts.

Notwithstanding these issues, it is relatively clear that cyber-attacks can be understood as processes that are caused to occur due to the actions of one or more agents. These agents are what we might call attackers, for they are agents who perpetrate an act that qualifies as a cyber-attack. Such agents may also be glossed as threats or (perhaps) threat agents. In particular, a threat agent will be an agent

that has the objective of causing harming to another agent (or agent collective). In the case of cyber attackers, this objective will be realized via their participation in processes that qualify as cyber-attacks.

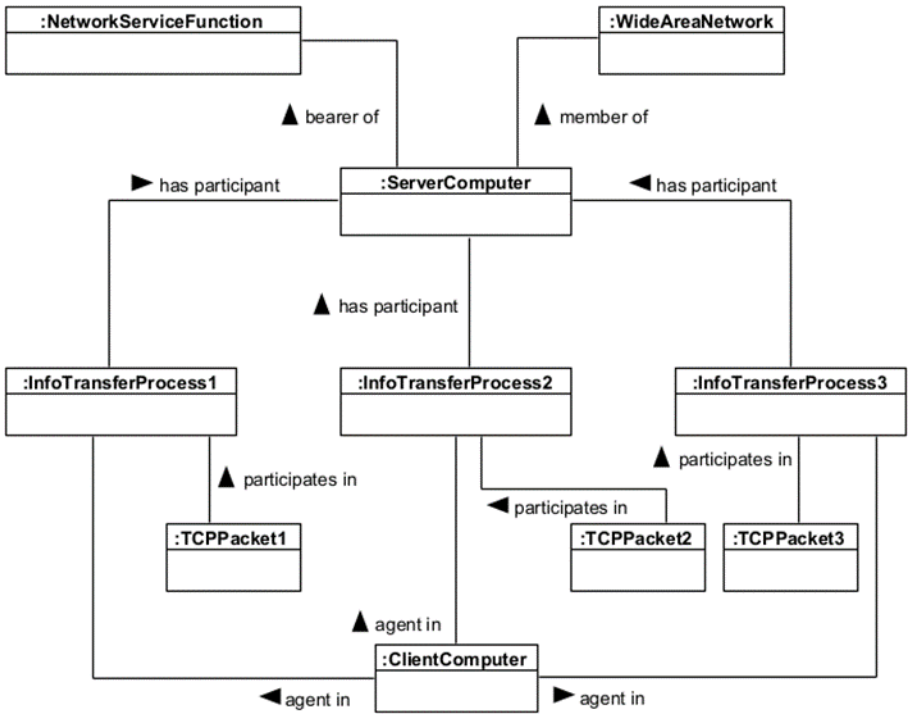

Figure 22. A SYN flood attack targeting a network server (adapted from Donohue et al. 2018).

What is it that makes something a cyber-attack as opposed to (let's say) a physical attack? This distinction is not made clear in the C3O, but it is reasonable to assume that the distinction relates to the material entities that are involved in the attack-related process. One possibility is that a cyber-attack can be understood as a cyber process (or computational process), specifically a process that involves (as participants) the use of computational equipment or other computational artefacts.

One example of a cyber-attack is depicted in Figure 22. This is an example of how a particular type of cyber-attack, specifically a SYN flood (or half-open attack), could be represented in a BFO-conformant manner. A SYN flood attack is a type of Denial of Service (DoS) attack (see Figure 21), which aims to make a server unavailable to legitimate traffic by consuming all available server resources. This is achieved via the repeated sending of initial connection (SYN) requests to the server without a corresponding acknowledgement of the server's response. In Figure 22, the server computer is represented by :SERVERCOMPUTER and the source of the SYN requests is represented by :CLIENTCOMPUTER (for the sake of simplicity, we have not sought to represent the human attacker in this scenario). The :CLIENTCOMPUTER is an ***agent in*** multiple information transfer processes, which involve the communication (or transfer) of Transmission Control Protocol (TCP) packets. The :SERVERCOMPUTER is depicted as being a member of a :WIDEAREANETWORK and the ***bearer of*** a :NETWORKSERVICEFUNCTION. The implementation of the SYN flood attack (which consists of a succession of information transfer process instances) may lead the :SERVERCOMPUTER to temporarily lose its :NETWORKSERVICEFUNCTION due to the fact that it is overwhelmed. This may

then have implications for other network users, as well as the individuals who operate the :SERVERCOMPUTER. For these individuals, the :NETWORKSERVICEFUNCTION is a valuable thing, because it enables them to do things that they want to do (i.e., to do things that satisfy their objectives). In this sense, the occurrence of the SYN flood attack culminates in a **NEGATIVE VALUE EFFECT** that entails the loss of a **VALUED REALIZABLE ENTITY** that ***inheres in*** a **MATERIAL ENTITY** that qualifies as an **ASSET**.

4.6 Capability

In contrast to the aforementioned concepts, capabilities are relatively easy to situate within a BFO-conformant ontology. This is because capabilities feature as one of the classes included in the CCO. They have also been the focus of sustained attention by the BFO community (Donohue et al. 2018; Hagedorn et al. 2019; Merrell et al. 2022). According to Hagedorn et al. (2019),

> Capabilities are treated as dispositions that are borne by continuants. For the purposes of the ontology, a capability is simply defined as a beneficial disposition of some continuant to successfully be able to participate in a process in some pre-specified way. This implies not just participation, but some quality of participation. Thus, the state of bearing capabilities enables successful participation in various processes. (p. 635)

Given this characterization, we can understand capabilities as dispositions that are borne by entities that qualify as (independent) continuants. Note that there are no claims here as to the precise nature of the capability bearer. Such bearers may be human individuals, as well as non-human entities, such as computers, cats, and IoT devices. In a security-related context, capabilities of interest could be those of an attacker or the user of a technology. It is also possible that one will want to characterize the capabilities of IoT devices and artefacts as part of a security modelling or risk assessment process.

4.7 Vulnerability

As with capabilities, vulnerabilities are typically understood to be dispositional properties. This is true in both a philosophical (e.g., McKitrick 2018) and applied ontological setting. In CCO, vulnerabilities are represented as a particular type of dispositional property, called a **DISRUPTING DISPOSITION**:

> **DISRUPTING DISPOSITION** =$_{def.}$ A disposition the realization of which would disrupt a process some entity has an interest in.

VULNERABILITIES are then defined as follows:

> **VULNERABILITY** =$_{def.}$ A disrupting disposition the realization of which would disrupt a process that the bearer of the disrupting disposition has an interest in.

C3O extends the notion of vulnerability to include what are called **CYBER VULNERABILITIES**:

> **CYBER VULNERABILITY** =$_{def.}$ A vulnerability that is realized by a cyber-attack interfering with or destroying a device's ability to function normally.

In one sense, this can all seem rather straightforward, and we certainly do not wish to contest the idea that vulnerabilities are, at root, some sort of dispositional property. At the same time, however,

it ought to be noted that BFO is wedded to something called the *intrinsic disposition thesis*. According to this thesis, dispositions are grounded in qualities that are internal to the entity that is the bearer of the disposition. Thus, if X has vulnerability Y, then an exact duplicate of X should also have vulnerability Y; there ought to be nothing about the surrounding environment that is germane to X's possession of Y.

The problem with the intrinsic disposition thesis is that it doesn't seem to tally with everyday intuitions about vulnerability. Consider the following case, which is attributable to McKitrick (2018):

> A military target, a city, is protected by a Star Wars-like defense system. The system has sensors that bring out defenses when there is a threat, rendering the city invulnerable. However, the sensors and anti-aircraft weapons are all located outside the borders of the city and are built, maintained, and staffed by a foreign country. Should the defense system be disabled, the city would change from being invulnerable to being vulnerable. However, the city might remain intrinsically the same in all ways that are relevant to its vulnerability. (McKitrick 2018, chap. 8)

What McKitrick is referring to here is the possibility of so-called *extrinsic dispositions*. These are dispositional properties that inhere in an entity (in this case, a city), but the things that make it true that the entity possesses (or bears) the relevant dispositional property are not things that lie internal to the thing that is the bearer of the relevant disposition.

This raises a question regarding the ontic status of vulnerabilities as dispositional properties within BFO. BFO is clearly committed to the intrinsic disposition thesis, and this implies that vulnerabilities (as dispositions) should conform to this thesis. The problem, however, is that there are multiple cases in which a vulnerability does not conform to the intrinsic disposition thesis. The aforementioned city case is one example. For a human case, consider the idea that a woman (call her Joan) could be vulnerable to physical attack if she walks alone in a park at night. If, however, she is surrounded by a cadre of armed bodyguards, then it is hard to see why we would regard her as particularly vulnerable. The point here is that Joan is the same material entity across both these scenarios, the only thing that need change is the features of the environment in which Joan is situated. Accordingly, Joan's vulnerability must be an externally-grounded (extrinsic) disposition, not an internally-grounded (intrinsic) disposition. Because BFO does not recognize the existence of extrinsic dispositions, we are unable to represent the fact that Joan is vulnerable if she would walk alone. If she is not vulnerable when surrounded by bodyguards, then she cannot qualify as vulnerable when she is by herself (or indeed, in any other situation). The upshot is a dilemma: either the proponents of BFO are wrong to embrace the intrinsic disposition thesis, or they are wrong to assert that vulnerabilities (or at least all vulnerabilities) ought to be understood as dispositional properties. As far as we can tell, there is no way of circumventing this dilemma.

4.8 Security Mechanism

Within the security domain, there are multiple views as to the meaning of the term "security mechanism." Jawar et al. (2022) suggest that the security mechanism concept refers to the practices that protect (IoT) devices from attack. This suggests that security mechanisms ought to be regarded as occurrent entities, for practices are most readily understood as processes of one sort or another.

Mazzaquatro et al. (2018) understand security mechanisms to be general entities that mitigate vulnerabilities and protect some asset. Unfortunately, this doesn't tell us anything about the ontic

nature of security mechanisms. In particular, it doesn't tell us what sort of entity a security mechanism is.

Oliveira et al. (2022) offer the following definition of a security mechanism:

> A Security Mechanism is an object, which may be a simple physical object like a wall, a high-tech air defense system like the Israeli Iron Dome, an Agent like a policeman, a social entity like a security standard or anti-COVID-19 rules, that bears dispositions called Control Capability. (Oliveira et al. 2022, pp. 373–374)

This definition suggests that security mechanisms are objects. In particular, they are objects that bear particular capabilities. This definition is useful, for it makes it clear that security mechanisms are being understood as objects. What is more, the reference to both objects and dispositions helps us situate security mechanisms within the BFO framework. We can thus specify that security mechanisms are **`OBJECTS`** that are ***`bearers of`*** a particular sort of **`DISPOSITION`**, namely, a capability (see Section 4.6).

Unfortunately, there are some problems with the way that Oliveira et al. (2022) define security mechanisms. The first relates to the suggestion that (e.g.) anti-COVID-19 rules are a *bona fide* example of a security mechanism. The problem here is that it is difficult to see how something like a rule could exist as an object. To be sure, it may be the case that a rule is serialized as a body of text and presented to the Prime Minister on a sheet of paper. The sheet of paper would, in this case, qualify as an object. But the paper is not the same as the rule. In BFO, rules and regulations belong to the realm of generically dependent continuants, and these entities are disjoint from the realm of independent continuants, which is where we find things like objects and, according to Oliveira et al. (2022), security mechanisms. It is also hard to see how a mere rule, or regulation, or security standard could, by itself, be said to have a disposition to do anything. In what sense, exactly, does an anti-COVID-19 rule bear a control capability, or indeed any capability? It is not immediately obvious to us that a rule is disposed to do anything. It may participate in processes that lead to the acquisition of a capability or the manifestation of a capability, but these capabilities are likely to belong to something other than the rule itself.

Another problem relates to our wider understanding of the mechanism concept. Security mechanisms, we assume, must be a particular kind of mechanism. If so, then the definition of a security mechanism will need to adhere to our current best understanding of what mechanisms are. If security mechanisms fail to satisfy these criteria, then they cannot qualify as mechanisms.

Mechanisms have, in fact, been the subject of sustained philosophical enquiry in recent years. The mechanism concept has thus been studied as part of what is called neo-mechanical or mechanical philosophy (Glennan 2017; Glennan and Illari 2018). There have also been attempts to apply the mechanism concept to work in cyber-security (Spring and Hatleback 2017; Spring and Illari 2019), as well as other areas of computer science (Smart et al. 2020; Smart, O'Hara, and Hall 2021). While there are disagreements as to the precise meaning of the term "mechanism," the central features of the mechanism concept are captured in the following succinct definition by Illari and Williamson (2012):

> A mechanism for a phenomenon consists of entities and activities organized in such a way that they are responsible for the phenomenon. (Illari and Williamson 2012, p. 120)

As is clear from this definition, the building blocks of mechanisms are what are called "entities' and "activities." Entities are typically understood to be material objects, along with their associated properties. They are the physical parts of mechanisms—the things that make up the mechanism. Activities, by contrast, are typically conceptualized as the "producers of change" (Machamer, Darden, and Craver 2000, p. 3) and as the "causal components of mechanisms" (Craver 2007, p. 6). They "are the things that the entities do" (Craver and Darden 2013, p. 16). In many cases, the activities describe the nature of the interactions between the entities that make up a mechanism, as when we say that an enzyme (entity) phosphorylates (activity) a protein (entity), a neuron (entity) releases (activity) a neurotransmitter (entity), a human agent (entity) edits (activity) a Wikipedia entry (entity), and a human agent (entity) tags (activity) an online image (entity).

The definition by Illari and Williamson has been widely accepted by the philosophical community; accordingly, we will assume that this definition is broadly correct in telling us what it means for something to count as a mechanism. Given this, we can return to the definition by Oliveira et al. (2022) to assess its compatibility with the mechanism concept. Security mechanisms, recall, are being understood as a particular sort of mechanism. In view of this, the definition proposed by Oliveira et al. (2022) ought to conform to the more general definition proposed by Illari and Williamson (2012).

Unfortunately, the conformance is, at best, partial in nature. Oliveira et al. suggest that mechanisms are objects (i.e., material entities). Given the definition by Illari and Williamson, however, this does not seem entirely correct. A mechanism is not just an object; it is more a multiplicity of objects engaged in activities. Activities, however, are occurrent entities, which means that we cannot make sense of the security mechanism concept without a reference to occurrents. The problem for Oliveira et al. is that their definition does not refer to occurrents. Their focus is on objects and the dispositions of those objects. But an object + disposition is not a process, and it seems that processes are of central significance to our basic understanding of mechanisms. In particular, it doesn't appear appropriate to say that something could exist as a mechanism if there was not some sort of process to accompany the mechanism. Consider the first part of Illari and Williamson's definition: a mechanism, they suggest, is *for a phenomenon*. A phenomenon, in this case, is (*inter alia*) an occurrent entity, which is to say it is an event, a process, or a state. Given this, it is somewhat hard to see how a mechanism could exist in the absence of occurrents. If security mechanisms are existentially dependent on occurrents, then they cannot be mere objects. They must be something else.

Perhaps, then, security mechanisms are best understood as occurrent entities, as per Jawar et al.'s (2022) appeal to practices. Unfortunately, that doesn't seem correct either, for mechanisms are more than just occurrents; they are more akin to an amalgamation of both objects and occurrents.

Where does that leave us? As should be clear by now, the concept of a security mechanism poses a significant challenge to the designers of security ontologies. Mechanisms, it seems, cannot be understood as either one thing or the other, they are more a combination of different things. From an ontological perspective, it may be best to model mechanisms as objects that are disposed to do certain things, which is the approach adopted by Oliveira et al. (2022). While this is not consistent with the philosophical understanding of mechanisms, it does establish a point of contact with the way we tend to talk about mechanisms in the vernacular. We might, for example, speak of an automobile engine as the mechanism for a car's propulsion, even though the car is locked in the garage and the engine is inactive. From a neo-mechanical perspective, what we are referring to here is not so much a mechanism as it is the material entity (the engine) that is poised to trigger the instantiation of a

mechanism once the key is inserted into the ignition. From a purely pragmatic standpoint, then, we suggest that security mechanisms are represented along the lines of Oliveira et al. (2022) (i.e., as objects that bear certain sorts of dispositions). In future work, it will be important to assess whether this sort of ontological strategy can be reconciled with the burgeoning literature on mechanisms, as well as mechanistically related terms, such as mechanistic explanation (e.g., Craver and Tabery 2016).

Aside from pragmatic constraints, there is a further reason why the object + disposition view might be relevant to our understanding of mechanisms. This stems from the way in which security mechanisms are deemed to block, prevent, or counter the risk associated with a cyber-attack (e.g., Baratella et al. 2022). This establishes a potential point of contact with the notion of blocking dispositions, complementary dispositions, and protective resistance, all of which have been explored in a bio-medical context (Goldfain, Smith, and Cowell 2010, 2011).[24] The notion of a blocking disposition seems to be particularly relevant here. Consider that if one is the recipient of a vaccine that protects one against an infectious disease, then the risk of one suffering from the disease is greatly diminished (perhaps to zero). From an ontological perspective, one's participation in a vaccination process allows one to acquire a resistance to a disease. This resistance is expressed in the form of a blocking disposition, which prevents (or blocks) an infectious entity (e.g., a virus) from participating in a process that reflects the realization/manifestation of the relevant disease. The 'mechanism', in this case, would be the object (or objects) that bear the blocking disposition, e.g., an individual's immune system.

Perhaps the same can be said for security mechanisms, such that the participation of an object in some sort of intervention (or, as Jarwar et al., call it a practice) leads it to acquire a disposition that blocks the realization/manifestation of another disposition, specifically that inhering in an entity that is causally involved in a cyber-attack process; e.g., a cybercriminal or (perhaps) a computer virus or rogue AI system.

4.9 Other Concepts

We have now discussed some of the key concepts in the SOfIoTS ontology described by Jarwar et al. (2022). The remaining concepts are security goal, criticality, and fault.

We have already discussed goals and objectives in an earlier section. A security goal is, we suggest, best understood as a particular sort of goal/objective, namely one that refers to the preservation of a valued entity. As discussed by Jarwar et al. (2022), there are multiple kinds of security goals. Examples include the likes of availability, resilience,[25] and safety. Different types of security goals will naturally refer to different types of entities or collections of entities. On the whole, however, there seems little reason to think that the earlier approach to representing goals and objectives would be inapplicable to security goals. As with the goals held by particular individuals, we can conceptualize security goals as representations that concretize objectives, where an objective is a information content entity that refers to (or prescribes) a state-of-affairs (e.g., a situation) that is desirable to one or more agents.

[24] This expands on a further form of correspondence that centres on the notion of risk (see Section 7.2). Just as one can be exposed to a security risk, it seems plausible that one could also be exposed to a health risk.
[25] See Daniel (2014), for an ontological account of resilience.

According to Jarwar et al. (2022), a "fault is a trigger, which may lead to a failure." In CCO, the notion of a fault is best understood in terms of the **DAMAGED STASIS** class. The **DAMAGED STASIS** class is a subclass of the **STASIS** class. It is defined as follows:

> **DAMAGED STASIS** =*def.* A Stasis of Specifically Dependent Continuant in which some Independent Continuant bears a Quality or Realizable Entity that has suffered impairment (i.e., a decrease or loss) due to a previous action or event such that the Independent Continuant is now of lesser value, usefulness, or functionality.

The final concept is criticality. According to Jarwar et al. (2022), "the criticality concept is similar to the capability concept, however, it is mostly used to represent a negative sense and is a synonym for 'Severity' or SeverityScale […]" This suggests that criticality is an evaluative notion, which is to say it is an evaluation or appraisal of the severity of a situation, process, or some other entity. From a BFO perspective, this makes criticality a type of **ESTIMATE INFORMATION CONTENT ENTITY**, which represents the informational results of an assessment or evaluation process. Given this, criticality values (or criticicality levels) can be represented in the usual way that (literal) values are represented in BFO-conformant ontologies (see CUBRC 2020b).

5 THE INTERNET OF THINGS

In this section, we provide an initial ontological characterization of IoT devices from a CCO perspective. To our knowledge, this is the first attempt to consider how IoT devices might be accommodated by the CCO. The closest approximation to the present effort is the C3O (see Section 3.5). The C3O, however, is a domain-level ontology that is oriented to the more generic realm of computational entities (e.g., computer networks and computational processes). For this reason, it doesn't not include terms that denote concepts of interest to both the IoT and cyber-physical domains. The present section is an attempt to address this gap.

5.1 IoT Devices

The first issue to consider is the nature of an IoT device. What is an IoT device, exactly?

In a general sense, an IoT device is a **MATERIAL ARTIFACT**, which places IoT devices in the general metaphysical category of independent continuants. What distinguishes IoT devices from other types of **MATERIAL ARTIFACT** remains a little unclear. Nevertheless, IoT devices are typically understood as artefacts that possess both computational (information processing) and communicative capabilities. From a CCO standpoint, this suggests that IoT devices lie at the intersection of two types of **MATERIAL ARTIFACT**. Specifically, they are likely to be material artifacts that qualify as both **INFORMATION PROCESS ARTIFACTS** and **COMMUNICATION INSTRUMENTS**:

> **INFORMATION PROCESSING ARTIFACT** =*def.* A Material Artifact that is designed to use algorithms to transform some Information Content Entity into another Information Content Entity.
>
> **COMMUNICATION INSTRUMENT** =*def.* A Material Artifact that is designed to facilitate communication between at least two entities.

Perhaps the best way of understanding IoT devices is to consider their functionality: the things they were designed to do. In this respect, IoT devices are likely to be devices that are individuated with respect to functional criteria. In particular, IoT devices are likely to possess (to be the bearers of) the following high-level functionalities:

> **COMMUNICATION ARTIFACT FUNCTION** =*def.* An Artifact Function that is realized in a process in which meaningful signs are conveyed from one entity to another.
>
> **COMPUTING ARTIFACT FUNCTION** =*def.* An Artifact Function that is realized by an artifact participating in a computation process.

The first of these functions captures the idea that an IoT device is designed to communicate with other entities, typically, via one or more computer networks. The second function captures the idea that IoT devices are, at root, computational devices—devices that are poised to participate in computational processes of one sort of another.[26]

In addition to these core functionalities, IoT devices are likely to be the bearers of other functions. Many IoT devices, for instance, are designed to sense, observe, or detect information from the

[26] For more, on the nature of computational processes, see Piccinini (2007; 2015; 2018).

physical environment. Such devices—typically referred to as **SENSORS**—may possess one or more of the following functions:

> **IMAGING ARTIFACT FUNCTION** =*def.* An Artifact Function that inheres in Artifacts that are designed to produce visual representations of entities.
>
> **SENSOR ARTIFACT FUNCTION** =*def.* An Artifact Function that is realized in processes wherein its bearer is used to produce an output signal which reliably corresponds to changes in the artifact's environment.
>
> **MEASUREMENT ARTIFACT FUNCTION** =*def.* An Artifact Function that is realized during events in which an Artifact is used to measure one or more features of a specified object or class of objects.

Other IoT devices are designed to effect some sort of change in the physical environment, typically for control purposes. Such devices—commonly referred to as **ACTUATORS**—are likely to be distinguished by a different set of functions. A Heating, Ventilation, and Air Conditioning (HVAC) system, for example, is likely to be distinguished via its possession of a ventilation function:

> **VENTILATION CONTROL ARTIFACT FUNCTION** =*def.* An Artifact Function that is realized in processes in which some Artifact is used to control the quality of air in some space.

As with all functions, the functions of IoT devices are realized in processes of one sort of another. The CCO processes that are likely to be of greatest relevance to IoT devices include **ACT OF COMMUNICATION**, **ACT OF MEASURING**, and **ACT OF OBSERVATION**. These processes cover the sensing and communicative functionalities of IoT devices, but the CCO is somewhat lacking when it comes to processual entities that reflect the realization of actuator functionalities.

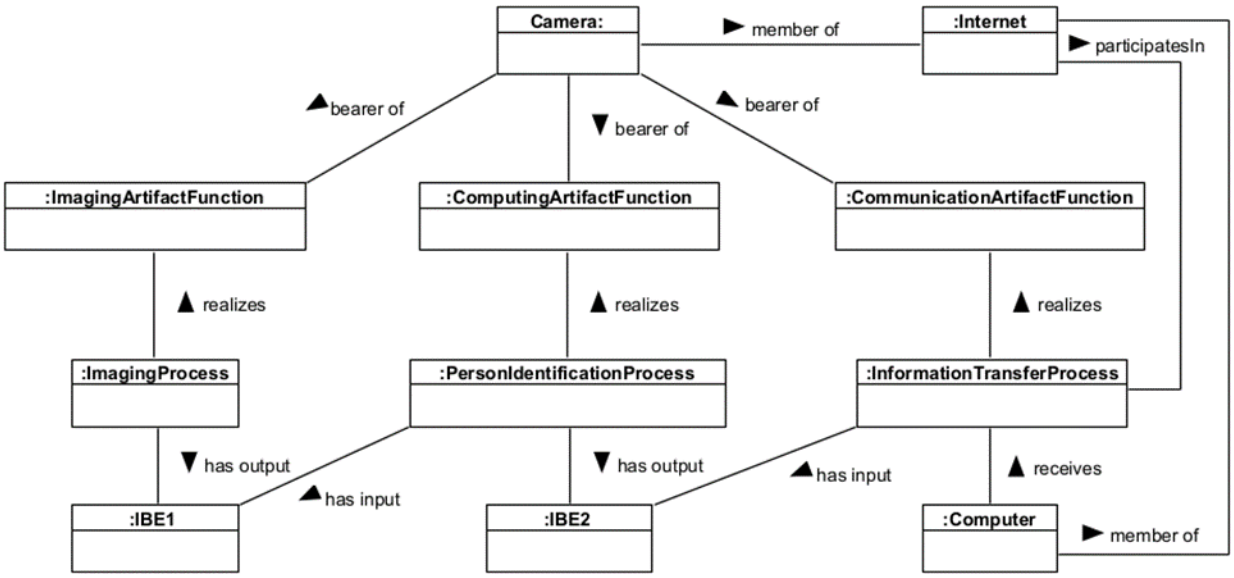

Figure 23. Device functionalities and their realization. The imaging, communication, and computational functions of an Internet-enabled imaging device are realized via the device's participation in multiple processes.

Figure 23 illustrates how the functionalities of an IoT device—in this case, an IoT-enabled CCTV camera—might be realized in a number of processes. The first thing to note here is that the CCTV camera is the bearer of multiple functionalities, specifically, an imaging artifact function, a computing artifact function, and a communication artifact function. These functions are realized in processes that capture visual data from the surrounding environment (`:IMAGINGPROCESS`), identify the object depicted in the visual data (`:PERSONIDENTIFICATIONPROCESS`), and communicate the identity-related information to another entity (`:INFORMATIONTRANSFERPROCESS`). As can be seen in Figure 23, processes can produce **`INFORMATION BEARING ENTITIES`** as output, which are then input to other processes. The `:INFORMATIONTRANSFERPROCESS` is a process that communicates information from one entity (`:CCTVCAMERA`) to another entity (`:COMPUTER`). This process is an instance of a new class—**`INFORMATION TRANSFER PROCESS`**—that is not included in the CCO. This class is defined as follows:

> **`INFORMATION TRANSFER PROCESS`** $=_{def.}$ A process in which an information bearing entity is communicated from one entity (the sender) to another entity (the receiver).

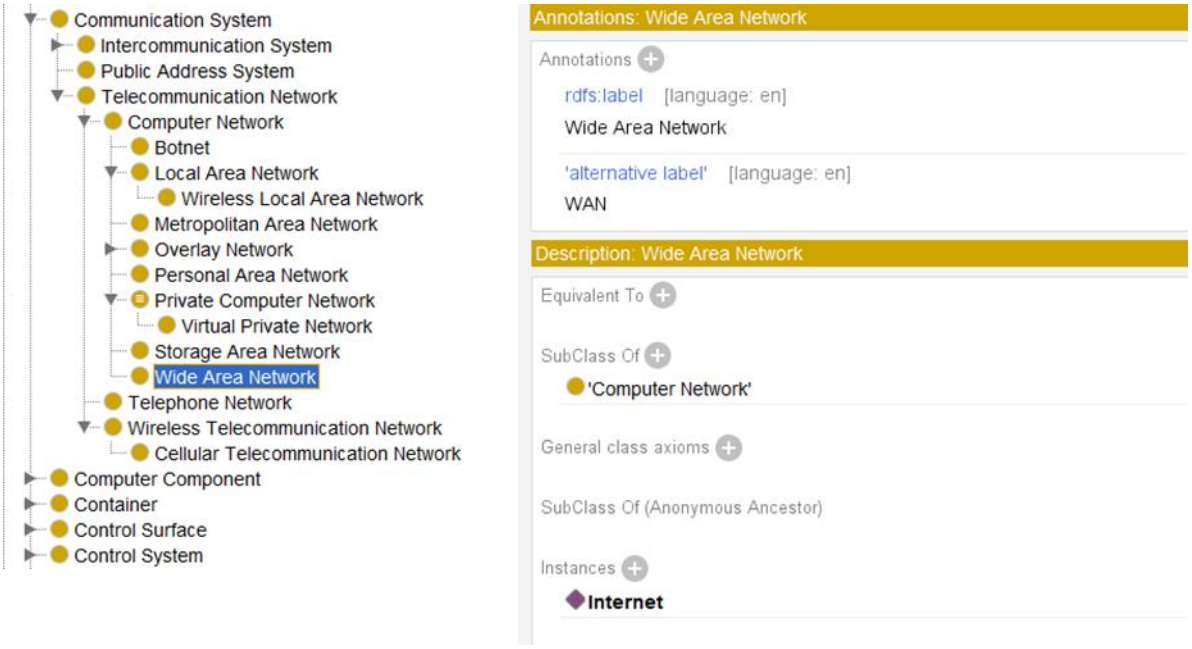

Figure 24. Computer Networks. The Internet is represented as an instance of Wide Area Network.

IoT devices typically communicate using various technologies like Wi-Fi, Bluetooth, Zigbee, or cellular networks, which allow them to connect to the Internet and exchange data with other devices or cloud-based platforms. This directs attention to the relationship between IoT devices (as **`MATERIAL ARTIFACTS`**) and computer networks. Following Donohue et al. (2018), we use the CCO ***`member of`*** relation to represent an IoT device's 'membership' of the Internet (see Figure 23). In Figure 23, the Internet is represented as an instance (or individual) as opposed to a class. This concurs with the C3O, which represents the Internet as an instance of the class **`WIDE AREA NETWORK`**, where a **`WIDE AREA NETWORK`** is defined as:

> **`WIDE AREA NETWORK`** $=_{def.}$ A Computer Network that is designed to connect several resources and is not restricted to a geographical location.

And **COMPUTER NETWORK** is defined as:

> **COMPUTER NETWORK** =*def.* A telecommunication network that is designed to allow the exchange of data between two or more computers connected to the network.

Figure 24 shows the position of these network-related classes in the C3O.

The effort to taxonomize IoT devices is complicated by the many varieties of IoT devices that are now available. The following highlights some of the high-level categories that might be used to taxonomize IoT devices:

- **Wearable Devices:** Devices that are worn or attached to the body, such as smartwatches, fitness trackers, or health monitoring devices.
    - Smartwatches
    - Fitness trackers
    - Health monitoring devices
    - Smart glasses
    - Smart clothing
- **Home Automation Devices:** Devices used for automating and controlling home systems, such as smart thermostats, smart lighting, or smart locks.
    - Smart thermostats
    - Smart lighting systems
    - Smart locks and security systems
    - Smart plugs and switches
    - Voice assistants (e.g., Amazon Echo, Google Home)
- **Industrial IoT Devices:** Devices used in industrial settings for monitoring and optimizing processes, such as machinery sensors, industrial control systems, or asset tracking devices.
    - Industrial sensors and actuators
    - Industrial control systems
    - Asset tracking devices
    - Predictive maintenance sensors
    - Connected machinery and equipment

- **Smart Energy Devices:** A device or appliance that incorporates advanced technologies and connectivity features to optimize energy usage, monitor consumption, and enable more efficient energy management.
    - Smart meters
    - Energy consumption monitors
    - Smart grid systems
    - Connected solar panels and energy storage systems
- **Connected Vehicles:** A vehicle that is equipped with advanced communication technologies and connectivity features that enable it to exchange data with external sources.
    - Connected cars
    - Telematics devices
    - Fleet management systems
    - Vehicle tracking and monitoring systems
- **Healthcare IoT Devices:** Devices used in biomedical and healthcare settings.
    - Remote patient monitoring devices
    - Medical wearables
    - Connected medical devices (e.g., insulin pumps, pacemakers)
    - Ambient assisted living systems
- **Smart Appliances:** Internet-connected household appliances that can be controlled and monitored remotely, such as smart refrigerators, ovens, or washing machines.
    - Smart refrigerators
    - Smart ovens
    - Smart washing machines
    - Smart dishwashers
    - Smart vacuum cleaners
- **Agricultural IoT Devices:** Devices used in agricultural settings.
    - Soil moisture sensors
    - Weather monitoring stations

  - Livestock tracking devices
  - Precision irrigation systems
  - Crop health monitoring systems

- **Environmental Monitoring Devices:** Devices that gather data from the physical environment, such as temperature, humidity, motion, light, or pressure sensors.
  - Air quality sensors
  - Water quality sensors
  - Weather stations
  - Noise level monitors
  - Pollution monitoring systems

- **Retail and Logistics IoT Devices:** A device used in the retail industry to enhance operational efficiency, improve customer experience, and gather data for analysis and insights.
  - Radio Frequency Identification (RFID) tags and readers
  - Inventory tracking systems
  - Smart shelves
  - Supply chain monitoring devices
  - Connected vending machines

As is perhaps clear from this list, IoT devices can be classified in multiple ways. IoT devices are sometimes classified according to the (industry) sector in which they are used (e.g., agricultural IoT devices). In other cases, IoT devices are classified using other criteria, such as their deployment in particular environments (Home Automation Devices, Smart Appliances), their status as a particular type of artefact (Connected Vehicles), and the sorts of entities on which they are deployed (e.g., Wearable Devices).

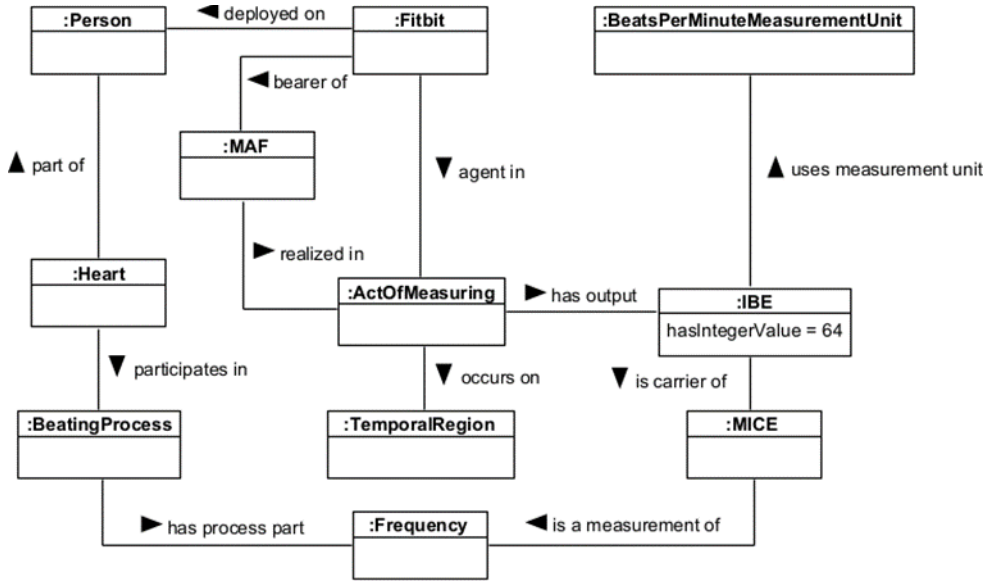

Figure 25. The measurement of a person's heart rate by a Fitbit device. [Acronyms: MICE (Measurement Information Content Entity), IBE (Information Bearing Entity), and MAF (Measurement Artifact Function).]

5.2 Sensors

As noted above, many IoT devices function as sensors, supporting the uptake and representation of information about some environmental property. Figure 25 depicts a scenario in which a Fitbit device is used to record information about an individual's heart rate. The first thing to note here is that we are using the ***deployed on*** relation as a shortcut way of specifying that the relevant device is located on the surface of the individual. A more detailed characterization of this relationship might feature an appeal to the various mereotopological relations that have been discussed in the BFO literature (e.g., Smith and Grenon 2004). For reasons of simplicity, we will not attempt to cover that literature here.

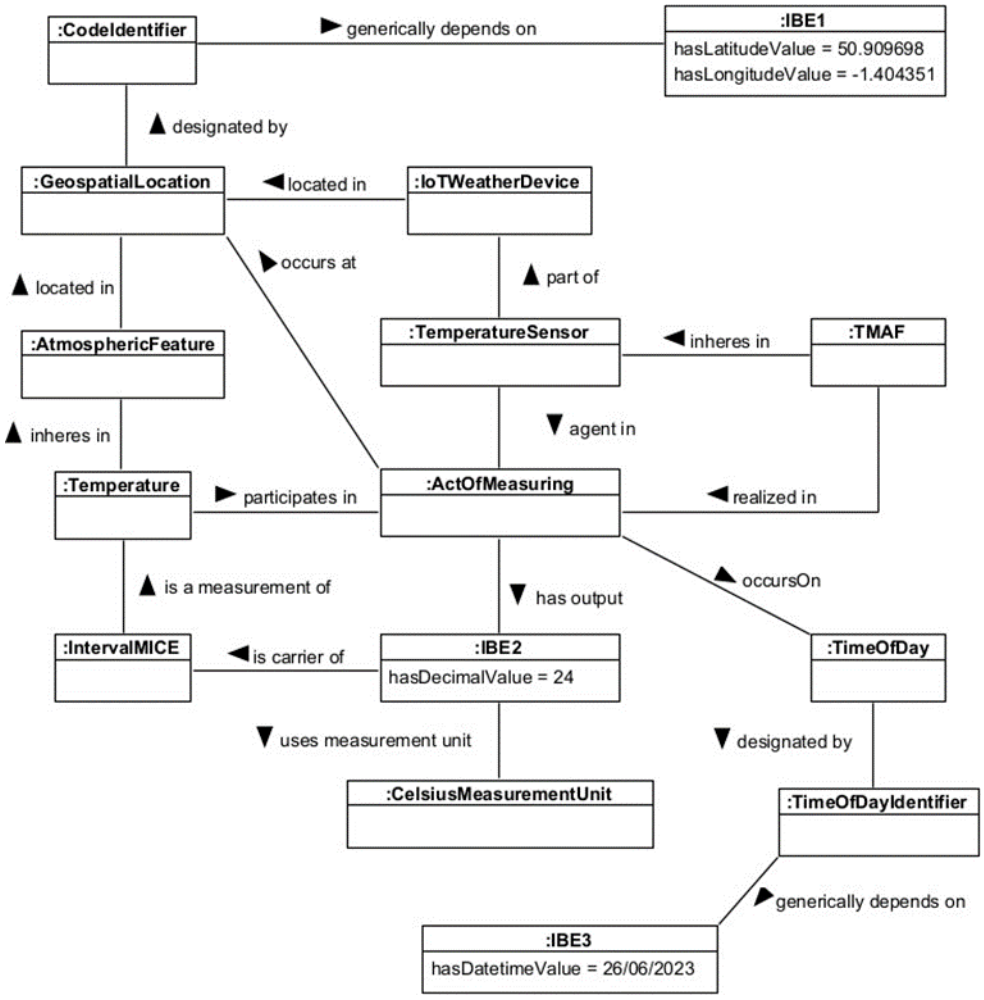

Figure 26. Temperature measurement by an IoT sensor. This example shows how information entities can be used to represent information pertaining to the time and location of measurements, as well as the value returned by an act of measuring. [Acronyms: IntervalMICE (Interval Measurement Information Content Entity), IBE (Information Bearing Entity), and TMAF (Temperature Measurement Artifact Function).]

In Figure 25, we have depicted a Fitbit device as the bearer of a **MEASUREMENT ARTIFACT FUNCTION**, which is realized in an **ACT OF MEASURING** process. The output of this process is an **INFORMATION BEARING ENTITY** (a digital object) that carries (is carrier of) information content (**MEASUREMENT INFORMATION CONTENT ENTITY**) pertaining to the beating frequency of a person's heart. In this case, `:FREQUENCY` is an instance of the **PROCESS PROFILE** class (this instance is a proper part of the heart's `:BEATINGPROCESS`).

The aim of Figure 25 is to show how the measurements made by an IoT device might be represented in a manner that conforms to the CCO. Figure 26 extends this example to show how information about environmental properties might be represented in CCO. Here, we see that a sensor (`:TEMPERATURESENSOR`) forms part of an IoT device (`:IOTWEATHERDEVICE`). The sensor is the bearer of a particular function (an instance of **TEMPERATURE MEASUREMENT ARTIFACT FUNCTION**), and this function is realized in a particular process, namely, (`:ACTOFMEASURING`). The

output of this process is an **INFORMATION BEARING ENTITY**,[27] which is the carrier of an **INTERVAL MEASUREMENT INFORMATION CONTENT ENTITY** that is about (***is a measurement of***) an environmental quality, specifically, the temperature of the ambient environment. The actual temperature reading is associated with the information bearing entity that is produced by the measurement process, which is the convention adopted by the CCO (see CUBRC 2020b). In addition, Figure 26 shows how measurements are tied to measurement units via the ***uses measurement unit*** relation.

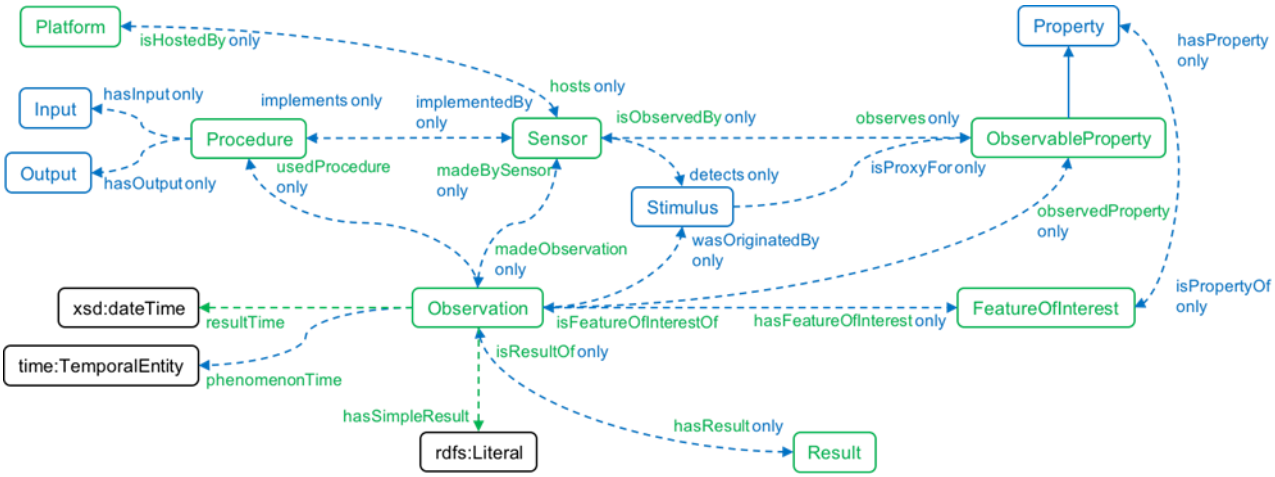

Figure 27. Classes and relations associated with observations in the SSN and SOSA ontologies. [Source: https://www.w3.org/TR/vocab-ssn/]

A common ontology used for sensors and sensor-related observations is the SSN ontology (Compton et al. 2012; Haller et al. 2019), which incorporates terms from other relevant ontologies such as the Sensor, Observation, Sample, and Actuator (SOSA) ontology. Figure 27 shows a subset of the classes and relations that are used to represent various aspects of sensor networks, including sensors, observations, features of interest, and the relationships between them. Table 4 describes how these classes map to terms in the BFO and CCO ontologies. While there are a number of differences between SSN and CCO, most of these differences are terminological in nature. Accordingly, there is no reason why SSN-based representations of IoT sensors and sensor-related information exchanges could not be rendered in a CCO-compliant format.

Table 4. Mapping of SSN sensor/observation terms to BFO/CCO.

SSN Term	BFO/CCO Term	Comment
FeatureOfInterest	**INDEPENDENT CONTINUANT**	SOSA features of interest are the entities whose property is being observed or measured.
Input	**CONTINUANT**	The inputs of a procedure/algorithm.
ObservableProperty	**QUALITY**	BFO qualities are (categorical) properties that inhere in independent continuants.

[27] Note that we allow both information content entities and information bearing entities to be the output of processes. In general, however, we assume information bearing entities to be the primary outputs of measurement-related processes.

SSN Term	BFO/CCO Term	Comment
Observation	**ACT OF OBSERVATION (OR MEASURING)**	SOSA procedures map to BFO processes.
Output	**CONTINUANT**	The outputs of a procedure/algorithm.
Platform	**MATERIAL ENTITY**[28]	We use the ***deployed on*** relation to represent the relation between sensors and sensor platforms.
Procedure	**ALGORITHM**	Algorithms are types of directive information content entities.
Property	**SPECIFICALLY DEPENDENT CONTINUANT**	Properties are specifically dependent continuants.
Result	**INFORMATION BEARING ENTITY**	The information bearing entity that is produced as a result of an Act of Observation/Measuring.
Sensor	**SENSOR**	Sensors are types of material entities.
Stimulus	**OCCURRENT**	The SSN stimulus class represents events (occurrents) that cause a sensor to participate in a process (e.g., an Act of Observation).

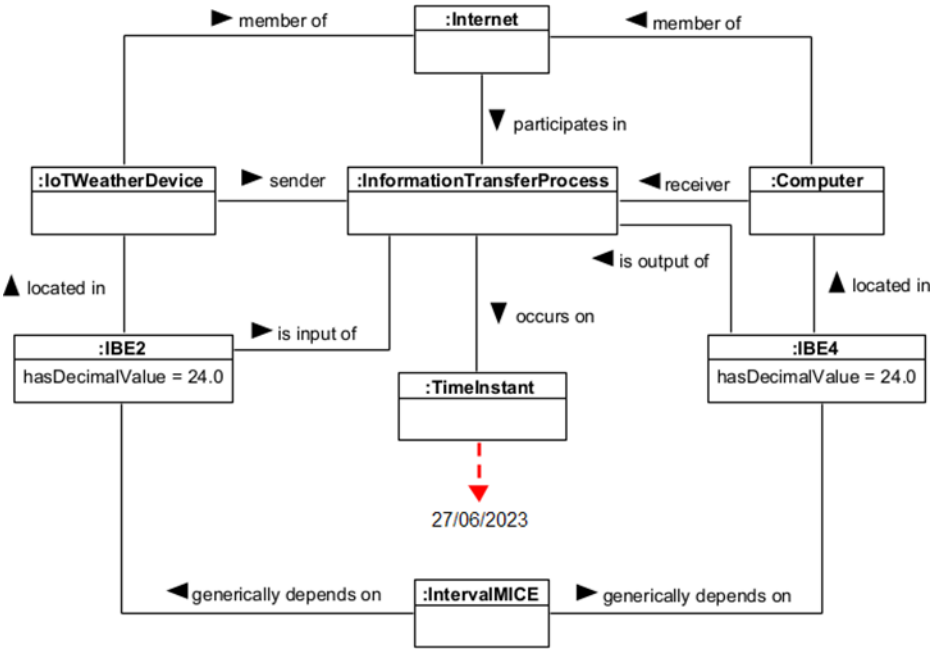

Figure 28. Transfer of measurement data from an IoT device to a central computer.

5.3 Communication

As noted in Section 5.1, communication is one of the core functionalities of IoT devices. Figure 28 extends the temperature measurement scenario depicted in Figure 26, with a view to showing how

[28] Note that CCO has a **SENSOR PLATFORM** class that is a specialization of **MATERIAL ENTITY**.

measurement data is communicated from one material artifact (:IOTWEATHERDEVICE) to another material artifact (:COMPUTER) via the :INTERNET. The actual communication occurs as part of :INFORMATIONTRANSFERPROCESS (see Section 5.1). The information bearing entity generated by :IOTWEATHERDEVICE serves as the input to the process, and the output of the process is another information bearing entity. This results in two information bearing entities, one of which is ***located in*** :IOTWEATHERDEVICE while the other is ***located in*** :COMPUTER. Despite this replication of information bearing entities, note that both these entities are associated with the same information content entity. Thus, while the carriers of information content (the information bearing entities) are distinct, the content of these entities is not.

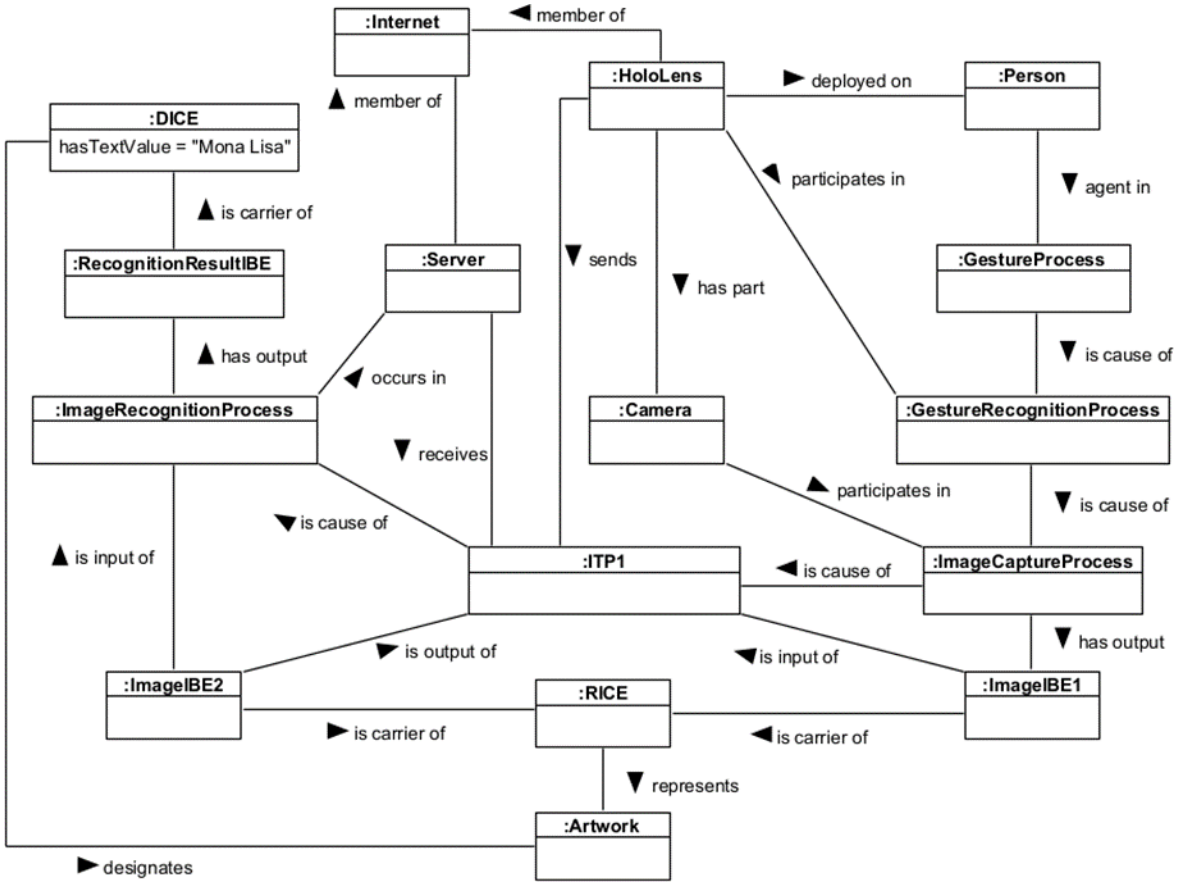

Figure 29. Recognition of artistic works using the HoloArt app. [Acronyms: RICE (Representational Information Content Entity), IBE (Information Bearing Entity), DICE (Designative Information Content Entity).]

5.4 Computation

While some IoT devices may do little more than transduce and transfer information, there is no reason why an IoT device cannot participate in computational processes that work to transform information. Figure 29 depicts a scenario in which a Microsoft HoloLens device is used to recognize works of art as part of an augmented reality application. This scenario is based on a real-world application, dubbed the HoloArt app (Smart 2021). The app enables human users to 'click on' physical works of works in the local environment. This gesture triggers a cascade of computational processing,

some of which occurs on the HoloLens device, while the rest occurs on a remotely-situated cloud computer. In Figure 29, the act of 'clicking on' an artwork is represented by the :GESTUREPROCESS. This triggers a :GESTURERECOGNITIONPROCESS, which in turn triggers an :IMAGECAPTUREPROCESS. In essence, the gesture causes the HoloLens device to take a picture of whatever it is the user is looking at. The result of the :IMAGECAPTUREPROCESS is an information bearing entity, which serves as the input to an information transfer process, which then culminates in the production of another information bearing entity on a server computer (:SERVER). This computer performs an image recognition process, which identifies the name of the artwork that lies within the human user's field of view.

5.5 Quality of Information

When it comes to issues of observation and measurement, it is often important to represent the accuracy or precision of the resultant information. In the case of the HoloArt app, for example, there are no guarantees that the outputs of the :IMAGERECOGNITIONPROCESS will be correct. In Figure 29, the output of the :IMAGERECOGNITIONPROCESS is an information bearing entity whose content is the name of the artwork (e.g., a painting) that the user wishes to identify. In the case of the HoloArt app, this information is communicated back to the HoloLens device and presented to the user. If the information is correct, then the user will be able to correctly identify the target artwork. If, however, the information is not correct, then the informational deliverances of the HoloLens do not amount to much; they are best a form of 'fake news'.

The HoloArt app contains (human factors) safeguards to prevent the possibility of a human user being led astray by incorrect results. More generally, however, we often want to represent the accuracy or veracity (or quality) of information produced by a sensing/computational process. CCO includes a number of classes that can be used to represent such information. For present purposes, the two most important classes are **RELIABILITY MEASUREMENT INFORMATION CONTENT ENTITY** and **VERACITY MEASUREMENT INFORMATION CONTENT ENTITY**. These are defined as follows:

> **RELIABILITY MEASUREMENT INFORMATION CONTENT ENTITY** $=_{def.}$ A Measurement Information Content Entity that is a measurement of the extent to which an entity can consistently produce an outcome.
>
> **VERACITY MEASUREMENT INFORMATION CONTENT ENTITY** $=_{def.}$ A Measurement Information Content Entity that is a measurement of the extent to which a description conforms to the reality it describes.

Figure 30 shows how the second of these two information content entities—i.e., **VERACITY MEASUREMENT INFORMATION CONTENT ENTITY**—could be used to represent estimates of the veracity of the information produced by the :IMAGERECOGNITIONPROCESS. The informational result of the :IMAGERECOGNITIONPROCESS is :RECOGNITIONRESULTIBE, which serves as the input to an estimation process (:ACTOFESTIMATION). This process produces another information bearing entity (:VERACITYIBE), which records the level of confidence in the veracity of the information produced by :IMAGERECOGNITIONPROCESS. An important feature of this example is the relationship between the two information content entities named :DICE and :VMICE. :DICE represents the content of the information carried by :RECOGNITIONRESULTIBE, while :VMICE

represents the content of the information carried by :VERACITYIBE. Both these information content entities have referents that are other entities. That is to say, both information content entities are about (***is about***) other entities. In the case of :DICE, the referent is a particular real-world object, namely, a work of art (this is established by the ***designates*** relation, which is a sub-property of ***is about***). In the case of :VMICE, however, the referent is just another information content entity, namely, :DICE. In essence, what this example shows is that we can use information content entities to capture metadata about other information content entities. In this case, we are attempting to represent a particular sort of metadata, namely, information about the veracity of information. There is, however, no reason why this approach cannot be generalized to report on other aspects of information quality.

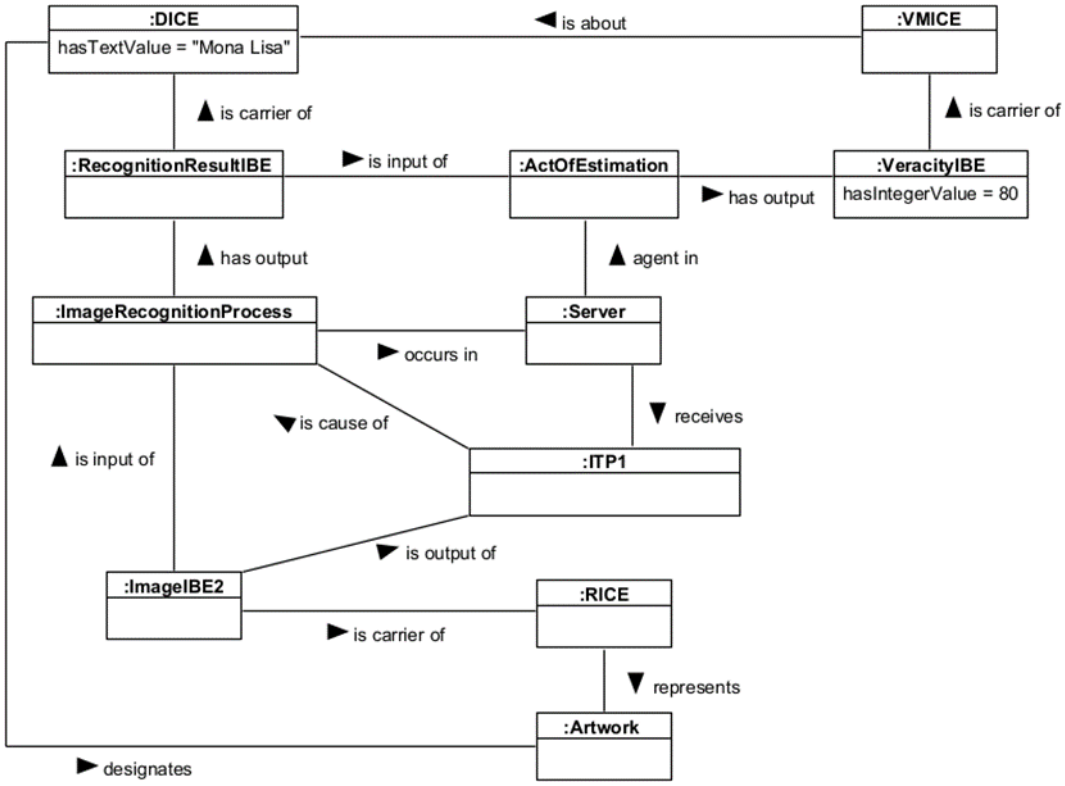

Figure 30. Estimation of veracity. [Acronyms: RICE (Representational Information Content Entity), IBE (Information Bearing Entity), DICE (Designative Information Content Entity), VMICE (Veracity Measurement Information Content Entity).]

5.6 Actuators

While the majority of IoT devices are designed to participate in some sort of sensing-related activity (e.g., a process that realizes a **SENSOR ARTIFACT FUNCTION**), some IoT devices can participate in actuation processes. In short, some IoT devices can effect changes in their environment by activating processes that exert a causal influence on these environments.

In Section 5.2, we provided an overview of the classes included in the SSN ontology, specifically, those that are intended to support the representation of sensor processes and sensor data. The actuation-related counterparts to these classes are depicted in Figure 31. Table 5 then reports the mapping

between these classes and the classes defined as part of the BFO and CCO ontologies. As can be seen from Table 5, the sosa:Actuator class has a direct mapping to the **ACTUATOR** class in CCO, where an actuator is defined as:

> **ACTUATOR** $=_{def.}$ A Transducer that is designed to convert some control signal into mechanical motion.

The mapping for other classes is, unfortunately, not so clear-cut. Two classes that look to be of particular importance are sosa:ActuatableProperty and sosa:FeatureOfInterest. The SSN ontology offers the following descriptions of these classes:

> **ActuatableProperty:** An actuatable quality (property, characteristic) of a FeatureOfInterest.

> **FeatureOfInterest:** The thing whose property is being estimated or calculated in the course of an Observation to arrive at a Result, or whose property is being manipulated by an Actuator, or which is being sampled or transformed in an act of Sampling.

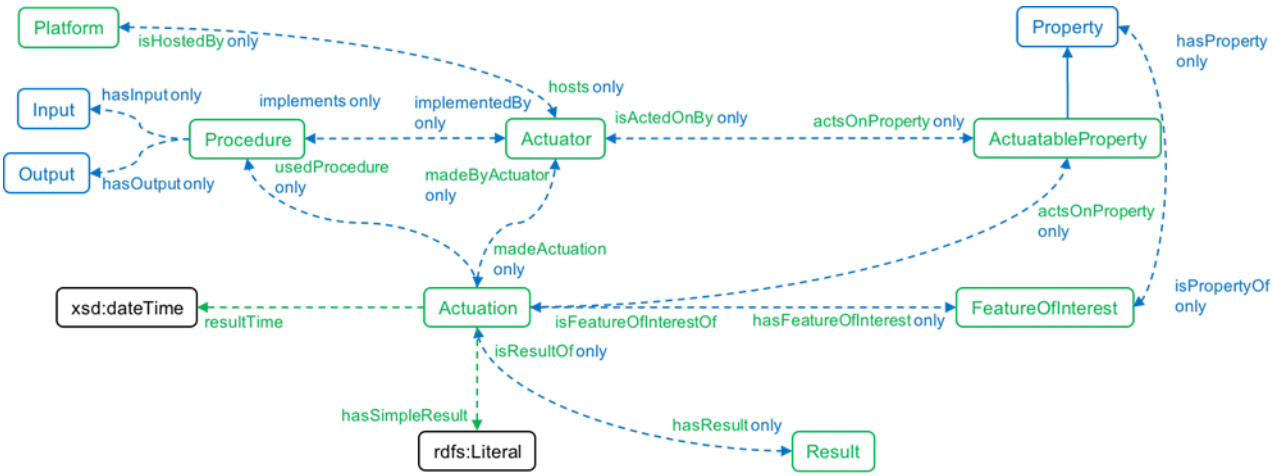

Figure 31. Classes and relationships associated with actuation in the SSN and SOSA ontologies. [Source: https://www.w3.org/TR/vocab-ssn/.]

Table 5. Mapping of SSN actuation terms to BFO/CCO.

SSN Term	BFO/CCO Term	Comment
ActuatableProperty	**QUALITY**	SOSA actuatable properties are the things that are affected by an actuation.
Actuation	**PROCESS**	
Actuator	**ACTUATOR**	
FeatureOfInterest	**INDEPENDENT CONTINUANT**	
Input	**CONTINUANT**	
Output	**CONTINUANT**	
Platform	**MATERIAL ENTITY**	
Procedure	**ALGORITHM**	
Property	**SPECIFICALLY DEPENDENT**	

SSN Term	BFO/CCO Term	Comment
	CONTINUANT	
Result	**INFORMATION BEARING ENTITY**	

In view of these descriptions, it seems appropriate to regard 'features of interest' as referring to entities that are the bearer of properties. In BFO, recall, properties are typically understood to be specifically dependent continuants (or, more generally, dependent continuants) (see Section 3.2). These (dependent) continuants inhere in independent continuants, which includes the likes of material entities. For this reason, it seems appropriate to map the sosa:FeatureOfInterest class to the BFO **INDEPENDENT CONTINUANT** class. An ActuatableProperty is then a property (dependent continuant) that inheres in an independent continuant. As with sensors (see Table 4), we suggest that these properties are best understood as **QUALITIES**, which are a particular type of **SPECIFICALLY DEPENDENT CONTINUANT**.

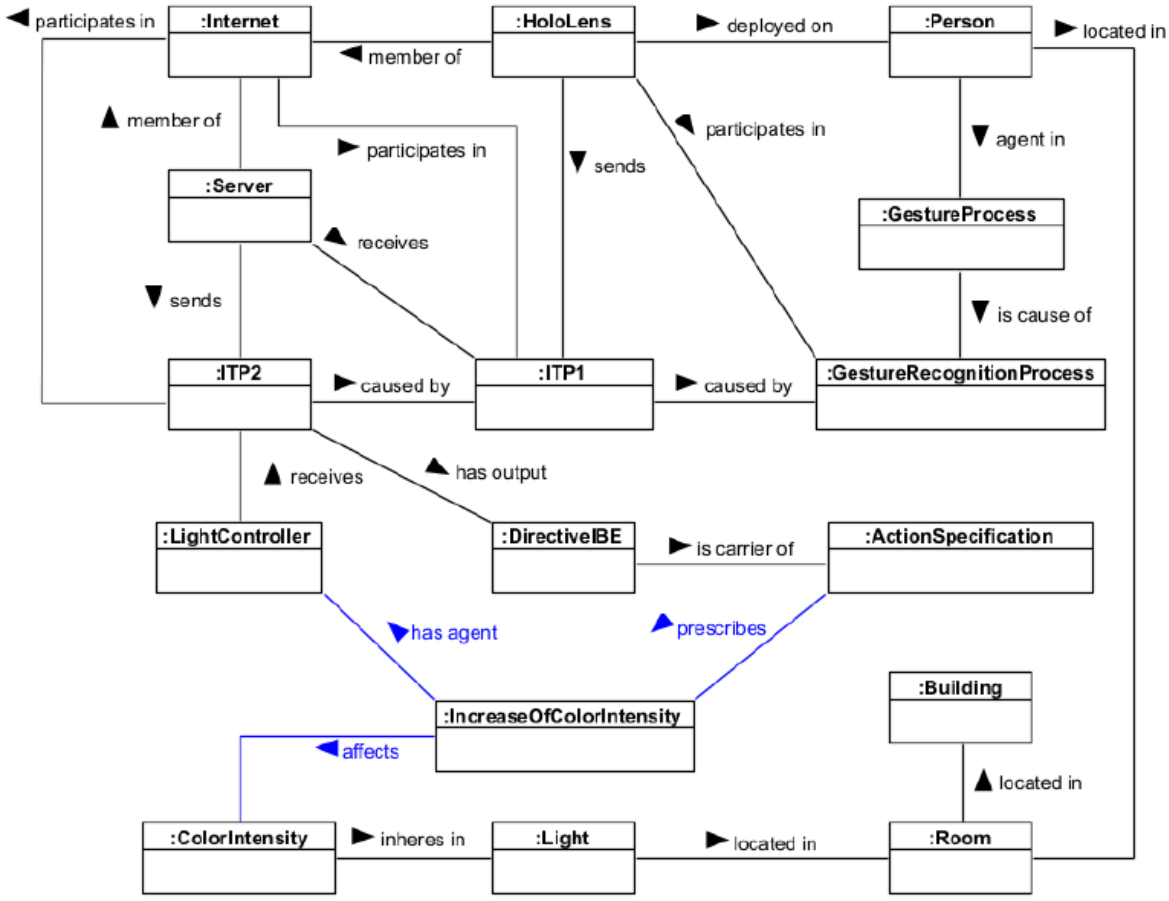

Figure 32. Use of one Internet-enabled device (a HoloLens) to control another (a light switch). [Acronyms: ITP (Information Transfer Process).]

Figure 32 shows how actuation processes can be represented in CCO. In particular, Figure 32 depicts a scenario in which a HoloLens device is used to adjust the lighting in a room. As with the HoloArt app case, we will assume that a human user participates in a gesture process, which is recognized by the HoloLens device. This triggers an information transfer process (:ITP1) to a remotely-situated computer (:SERVER), which then issues a command (via :ITP2) to an Internet-enabled light control device (:LIGHTCONTROLLER). The command, itself, is represented by :DIRECTIVEIBE, which is the carrier of information content (:ACTIONSPECIFICATION) pertaining to the type of action (or process) to be performed by the target actuator (i.e., :LIGHTCONTROLLER). As can be seen from Figure 32, the (:ACTIONSPECIFICATION) ***prescribes*** an :INCREASEOFCOLORINTENSITY, which is an instance of the **INCREASE OF QUALITY** class (i.e., a type of process). The :INCREASEOFCOLORINTENSITY process ***affects*** a quality, named :COLORINTENSITY, which ***inheres in*** a :LIGHT whose location corresponds to that of the individual who performed the original gesture.

5.7 Environmental Control

In Figure 32, it is a human user that triggers the cascade of causal processes that culminate in a technology-medicated shift (a **CHANGE**) of some environmental property. In some cases, however, the sensorimotor control loop may be closed, such that IoT sensors collect information that drives IoT actuators. Figure 33 presents the case of a HVAC control system, which is designed to control the temperature of a building. For the sake of simplicity, we have omitted the sensor-related components of this control process and limited the focus :TEMPERATURECONTROLPROCESS, which is a process that realizes the functionality of the :HVACSYSTEM by controlling air temperature. As with Figure 32, we are assuming that action-related commands are communicated to the :HVACSYSTEM via an information bearing entity (:DIRECTIVEIBE) that prescribes a particular course of action, specifically, an :INCREASEOFQUALITY process. The difference is that this command is ***prescribed by*** a **PLAN SPECIFICATION** as opposed to an **ACTION SPECIFICATION**.[29] In practice, this makes little difference, for both **ACTION SPECIFICATIONS** and **PLAN SPECIFICATIONS** are types of **DIRECTIVE INFORMATION CONTENT ENTITIES**. **PLAN SPECIFICATIONS**, however, come with objectives that specify the states-of-affairs or, recalling the discussion in Section 4.2, the situation that is to be achieved as the result of implementing a plan. This could be used to accommodate scenarios where the **PLAN SPECIFICATION** simply ***prescribes*** a recommended course of action, while the actual course of action (the precise sequence of actions) is left to the IoT actuator.

[29] Note that the **ACTION SPECIFICATION** class is not included in the CCO.

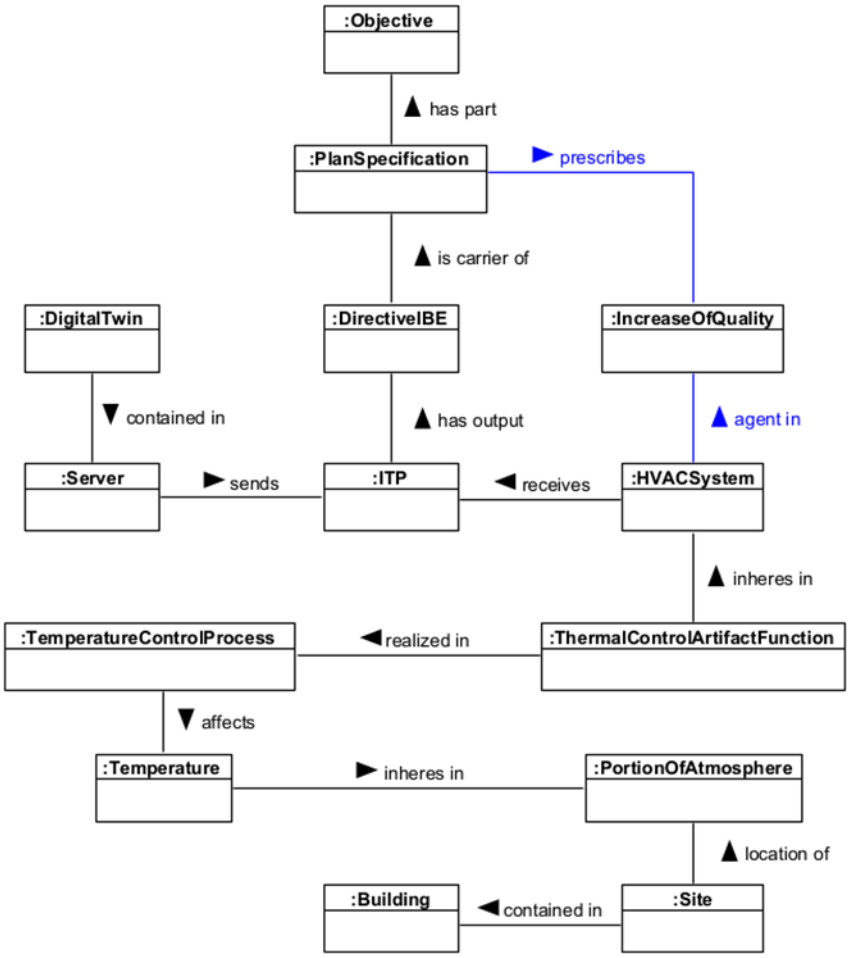

Figure 33. HVAC control system. [Acronyms: ITP (Information Transfer Process).]

A further feature of Figure 33 relates to the inclusion of a digital twin (`:DIGITALTWIN`). We will have more to say about digital twins in the next section. For present purposes, however, it is worth noting that digital twins are being classed as **`INFORMATION PROCESSING ARTIFACTS`**. This stems from the intuition that digital twins are a form of simulation software, and **`SIMULATION SOFTWARE`** is a type of **`INFORMATION PROCESSING ARTIFACT`**, at least according to the C3O.

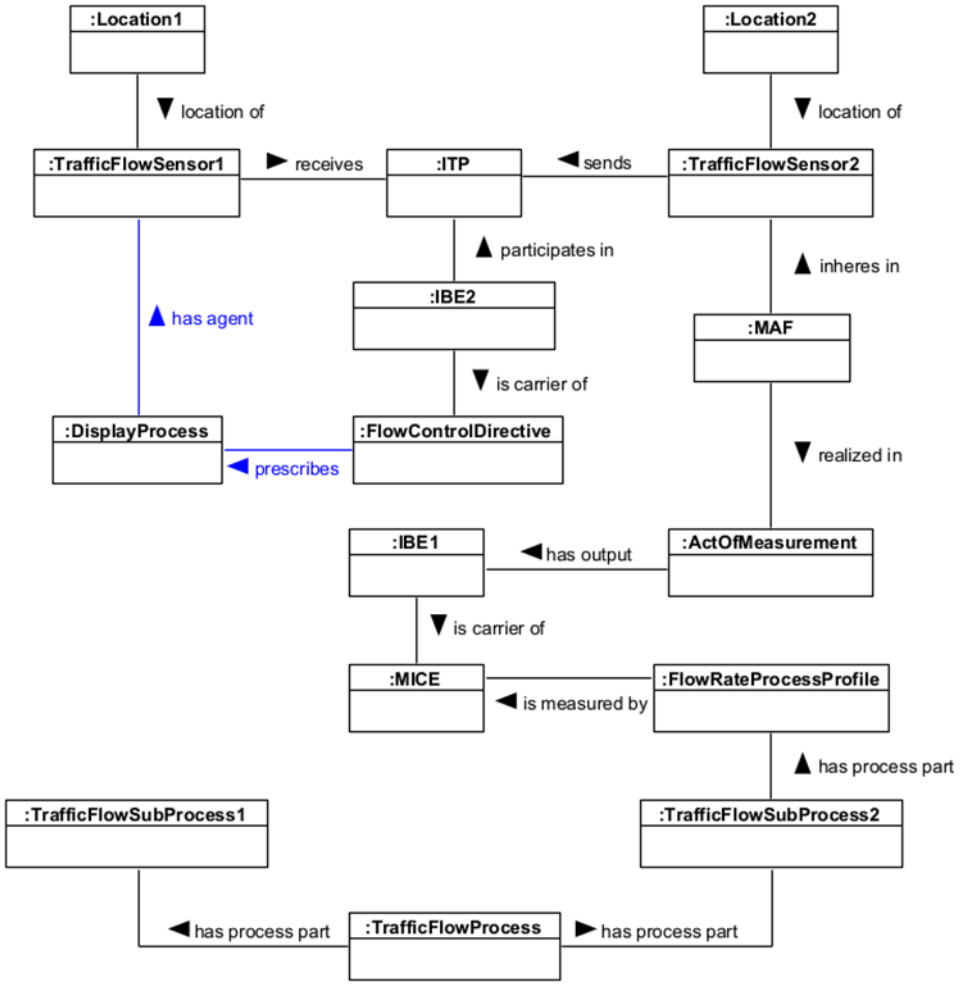

Figure 34. Traffic control system. [Acronyms: ITP (Information Transfer Process), IBE (Information Bearing Entity), MICE (Measurement Information Content Entity), MAF (Measurement Artifact Function).]

As a final example of the attempt to model the flow of information and control between IoT devices, consider Figure 34, which illustrates a case where one traffic flow sensor (`:TRAFFICFLOWSENSOR2`) communicates information to another traffic flow sensor (`:TRAFFICFLOWSENSOR1`) as a means of modulating a traffic flow process. In this scenario, `:TRAFFICFLOWSENSOR2` is located 'downstream' of `:TRAFFICFLOWSENSOR1`, so the traffic flow process (or sub-process) monitored by `:TRAFFICFLOWSENSOR1` has a bearing on the traffic flow process monitored by `:TRAFFICFLOWSENSOR2`. If `:TRAFFICFLOWSENSOR2` detects a decrease in the rate of traffic flow, it can issue an instruction to `:TRAFFICFLOWSENSOR1`, requesting that it participate in a `:DISPLAYPROCESS` that modifies the speed vehicles that participate in the larger `:TRAFFICFLOWPROCESS`.

5.8 Digital Twins

Figure 33 introduced the notion of a digital twin, which was cast as an **`INFORMATION PROCESSING ARTIFACT`**. Digital twins are artefacts that are intended to model the properties and dynamics of a real-world object, system, or process. They are created using real-time data from IoT devices and

other sources to model and simulate the behaviour, characteristics, and performance of a real-world (physical or material) entity. By capturing and analysing data from the target, real-world entity, a digital twin provides a detailed and dynamic replica that can be used for a variety of monitoring, analytic, and control-related purposes.

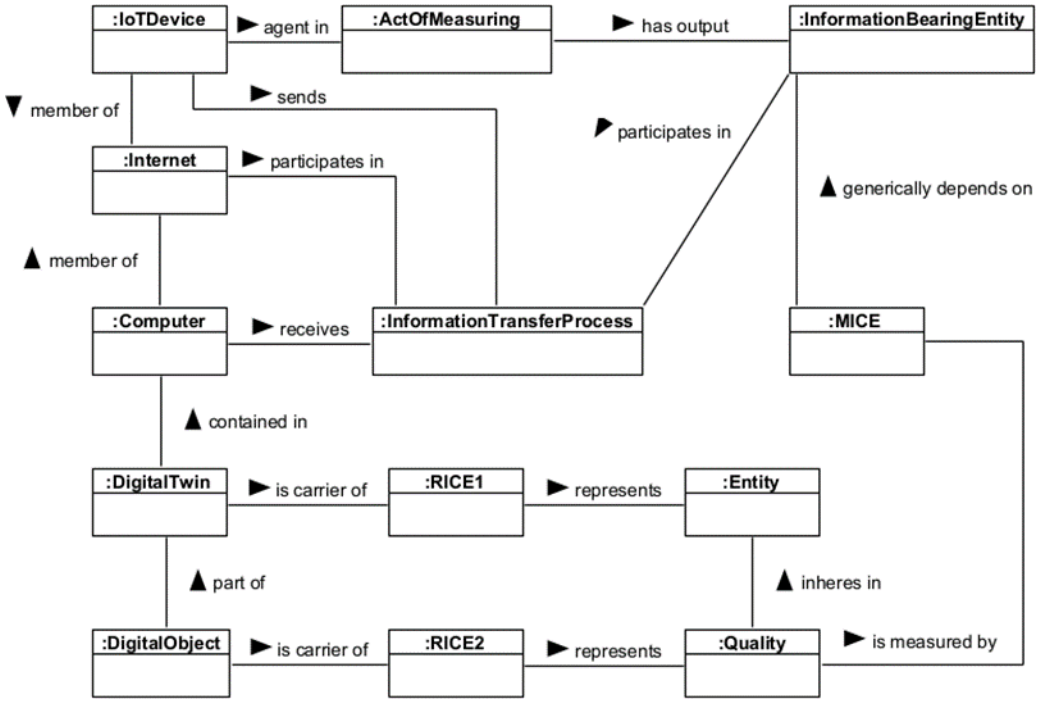

Figure 35. Representing digital twins in CCO. [Acronyms: RICE (Representational Information Content Entity), MICE (Measurement Information Content Entity).]

Figure 35 illustrates our approach to the representation of digital twins in the CCO. As noted above, digital twins are represented as information processing (or computational) artefacts. In Figure 35, the digital twin (`:DIGITALTWIN`) is contained in a conventional digital computer (`:COMPUTER`) that receives information from an `:IOTDEVICE` in the manner described above. Given their role in representing real-world entities, we suggest that digital twins are the carrier of a particular kind of information content entity, namely, a **`REPRESENTATIONAL INFORMATION CONTENT ENTITY`**. In CCO, representational information content entities are defined as information content entities that represent some entity. This entity could, of course, be anything, although we assume that it will mostly correspond to an **`INDEPENDENT CONTINUANT`**.

The digital twin in Figure 35 consists of a single object, named `:DIGITALOBJECT`.[30] This is intended to capture the idea that a digital twin consists of digital objects, each of which can be understood as an **`INFORMATION BEARING ARTIFACT`**. In C3O, for example, a digital object is defined as follows:

[30] In practice, of course, a digital twin will consist of multiple digital objects. For the sake of simplicity, however, only a single digital object is shown in Figure 35.

DIGITAL OBJECT $=_{def.}$ An Information Bearing Artifact that is designed to bear a collection of related values.

As with the larger :DIGITALTWIN in which the :DIGITALOBJECT is contained, :DIGITALOBJECT can be a carrier of information. In this case, :DIGITALOBJECT is the carrier of another **REPRESENTATIONAL INFORMATION CONTENT ENTITY**, which represents one of the qualities that inheres in the entity represented by :DIGITALTWIN. What we end up with, then, is the idea that a digital twin is an **INFORMATION PROCESSING ARTIFACT** that consists of a hierarchy of **DIGITAL OBJECTS**, each of which can be understood to represent a particular property or (perhaps a constituent object) of the **ENTITY** that is represented by the digital twin.

The final point to note is that the **QUALITY** represented by the aforementioned digital object is a quality that is measured by an IoT device as part of a specific measuring process, namely :ACTOFMEASURING. The information bearing entity produced by this measuring process is :INFORMATIONBEARINGENTITY, which participates in an information transfer process (:INFORMATIONTRANSFERPROCESS) via the :INTERNET. Figure 35 thus models a scenario in which an IoT device is used as a sensor to perform measurements of some target property (or **QUALITY**) for the purpose of updating a digital twin that represents an **ENTITY** that is the bearer of the target property. For the sake of brevity, we have not elected to model the reverse situation, where a digital twin is the source of commands that ***prescribe*** the behaviour of an IoT device. This process can, however, be understood as a relatively straightforward extension of the environmental control systems modelled in Sections 5.6 and 5.7. To the best of our knowledge, this is the first attempt to represent digital twins and IoT devices within the ontological framework of BFO and CCO.

6 HUMAN FACTORS

In earlier work, we identified human factors as one of the features that are missing from contemporary security ontologies (Jarwar, Tooth, and Watson 2022). This is especially true of ontologies that are developed in respect of IoT devices. The C3O, for example, includes classes representing common security concepts, but it does not feature classes representing human factors information. Likewise, the Internet of Things Security Ontology (IoTSEC), developed by Mozzaquatro and colleagues (Mozzaquatro, Jardim-Goncalves, and Agostinho 2015; Mozzaquatro et al. 2018), provides a useful taxonomy of assets, security mechanisms, threats, and vulnerabilities, but human factors seem to be somewhat under-represented in the ontology.

The omission of human factors looks to be important given the emphasis that is usually attached to human agents in the characterization of security-related risks. As noted by Oltramari et al. (2015):

> A fully predictive cyber security risk assessment model will take into account humans as risk factors, and as risk mitigators, and will enable the incorporation of metrics that go beyond the classic CIA [confidentiality, integrity, accessibility] vulnerabilities. […] As these arguments suggest, untangling the complexity of cyber security does not solely depend on pinning down the computational elements into play, but demands a thorough analysis of the human factors involved. (Oltramari et al. 2015, pp. 27–28)

In all likelihood, human factors issues are a common concern for *all* security ontologies. There are, however, reasons to think that these concerns are particularly salient when it comes to IoT devices and CPSs. Unlike a conventional computer, which might be shielded behind a firewall and used by a limited number of users, an IoT device operates in much more public space, and such devices are used by individuals with varying levels of expertise, knowledge, and skills.[31] What is more, such devices are sometimes embedded in a wider social and technological fabric, which makes it difficult to analyse risk and security issues in the absence of a consideration of the various forms of causal commerce that a given IoT device has with other entities.

Given this, let us accept the idea that human factors are an important element of any security ontology that is developed for the IoT domain. At this point, a couple of issues loom large. The first is what is a human factor, exactly, and what human factors are relevant to security-oriented analyses? This is a question about the nature of human factors and the identification of those factors that are important for security analysis. For the sake of convenience, let us call it the *analytic problem*.

A second issue relates to the ontological positioning of human factors vis-à-vis security ontologies. In short, where should we place the various ontological elements (classes and relations) that are intended to represent human factors information? Call this, the *positioning problem*.

One answer to the positioning problem is that the ontological elements ought to be included in the security ontology. The problem with this response is that it risks violating a design principle that is widely adopted in the BFO community (Hagedorn et al. 2019). This principle is one of a modular approach to ontology development. In particular, the idea is that rather than develop a single

[31] There have also been changes in working practices since the COVID-19 pandemic, with more people now opting to work from home. This poses a challenge to traditional (enterprise-centric) security models, which tend to assume that individuals will be working within a particular location. We are grateful to George Miguel at Cambiont.com for bringing this issue to our attention.

monolithic ontology that incorporates every entity of interest in a particular domain, we ought instead to develop a suite of ontologies that focus on particular kinds of entities. These ontologies can then be combined for a multiplicity of different purposes. Consider that a human factors ontology might be of general relevance to multiple sorts of application, not all of which are concerned with matters of security. The best approach here is to develop ontologies that focus specifically on human factors. These ontologies can then be used as part of a (perhaps more specific or fine-grained) effort to develop security ontologies where a subset of human factors are deemed to be important. Given this, it would clearly be a mistake to expect human factors classes to be asserted in an ontology for IoT-related security. Rather than those these classes being asserted within the security ontology, it would be far more appropriate for the security ontology to import these classes from a wider nexus of ontologies. Ideally, then, we ought not to be thinking of a security ontology that 'includes' human factors, in the sense of asserting and defining these classes. Instead, we ought to be thinking of ontologies that 'reference' human factors via the use of ontology import relations and ontology mapping mechanisms.

At this point, the positioning problem looks to be resolved. Time, then, to turn our attention to the analytic problem. What sort of entities fall under the heading of a human factor, and which of these entities ought to be imported or referenced by an IoT security ontology?

The answer to this question is not straightforward. Within the academic literature, it is common to find discussions about the supposed importance of human factors and socio-technical issues to our understanding of security. In most cases, however, there is very little discussion of what those issues are and what their ontic character might be.

At the most general level, the term "human factors" refers to the properties possessed by human individuals and the processes in which these individuals participate. In response to the question "what are human factors?" ChatGPT yielded the following response:

> Human factors, also known as ergonomics, refer to the scientific discipline that studies the interactions between humans and their surrounding environment, products, systems, or processes. It focuses on understanding human capabilities, limitations, and behaviors to design and optimize systems that enhance human performance, safety, and well-being.

It seems, then, that we are dealing with the properties, features, or attributes of human individuals, specifically those that support their participation in processes that involve some sort of interaction with an extra-individual entity (e.g., an IoT device). From an ontological standpoint, human individuals are typically represented as agents or organisms, which puts them in the metaphysical category of material entities and thus independent continuants. The properties of these individuals are specifically dependent continuants, which will assume the form of either qualities or realizable entities. These specifically dependent continuants will inhere in particular human individuals, which will be regarded as the bearer of those specifically dependent continuants. Within the general category of specifically dependent continuants, we have realizable entities, which includes things like roles and dispositions. These realizable entities are properties whose actualization (or realization) depends on the instantiation of particular processes. Processes, in turn, are represented as occurrent entities in BFO. This speaks to the idea that human factors is a discipline that is concerned with both the characteristics of human individuals and the behaviours in which those individuals participate.

In BFO, the former would be represented as specifically dependent continuants, while the latter would be represented as processes.

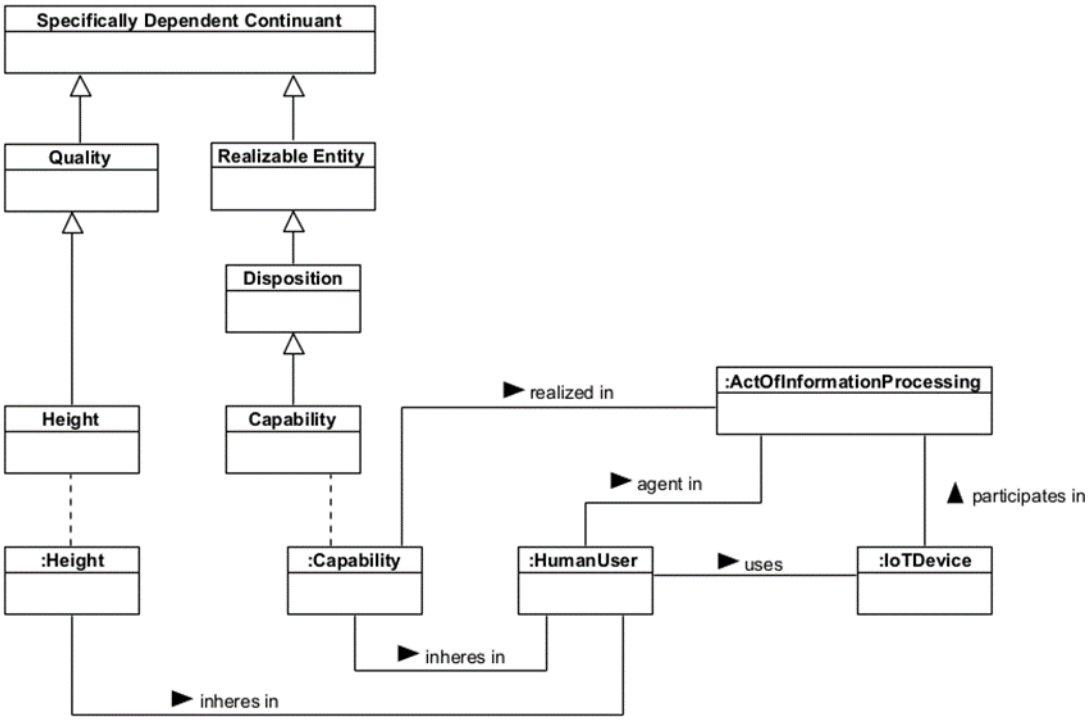

Figure 36. The capabilities and qualities of a human individual.

As a means of helping us understand how human factors might be represented in an ontology, let us consider a couple of examples. Figure 36 shows how the properties of a human individual would be represented in BFO. Note that there are two kinds of properties here. The first is the height of the individual, which in BFO is represented as a quality. The second is a capability, which is represented as a particular type of dispositional property and thus a realizable entity (see Merrell et al. 2022). Both these properties inhere in a particular human individual. This individual is represented as a participant in (or, as depicted here, an ***agent in***) a particular process. This process also features the participation of an IoT device, which is used by the individual as part of their performance of the process. In performing the process, the individual is said to be manifesting (or realizing) their capability. Accordingly, the process represents a realization of the capability that inheres in the individual.

Capabilities look to be particularly important when it comes to human factors considerations. In BFO, capabilities are typically understood as a particular kind of dispositional property. They are, at least, represented as a type of realizable entity, which are the sorts of entities that depend on processes for their actualization (or realization). According to Merrell et al. (2022), a capability is a type of dispositional property that is related to the notion of a function. In particular, they suggest that capabilities are a class of entities that are intermediate between dispositions and functions:

> A disposition inheres in a material entity and is realized in a certain kind of process. An example is the disposition of a glass to break when struck, which is realized when it shatters. A function is a disposition which is (simply put) the rationale for the existence

> of its bearer. […] Capabilities are a special sort of disposition in that, like functions, they can be evaluated on the basis of how well they are realized. They differ from functions in that their realizations are not the rationale—not the primary reason—for the existence of their bearers. […] All functions are capabilities on the view we defend, but not all capabilities are functions. (Merrell et al. 2022, p. 1)

We will have more to say about the status of capabilities as dispositional properties below. For the time being, however, let us accept that capabilities are at least one of the things that ought to be accommodated by an ontology that is concerned with human factors. In BFO, these capabilities will be represented as dispositions that are realized in processes, which is the state of affairs depicted in Figure 36.

While Figure 36 shows us how to represent the realization of a specific capability within a specific process, it is not particularly useful for the sorts of situations in which human factors considerations tend to surface. Consider that human factors tend to feature as part of the design process for some artefact or system, specifically one that will be used by human individuals. Figure 36 depicts a situation in which a human user is *already* interacting with a device (or has done so in the past). From a design perspective, however, what we want to know is not how a user *does* interact with an as yet non-existent device. Instead, what we want to know is how the design of a device ought to be informed by the capabilities that of a target user (or user community).

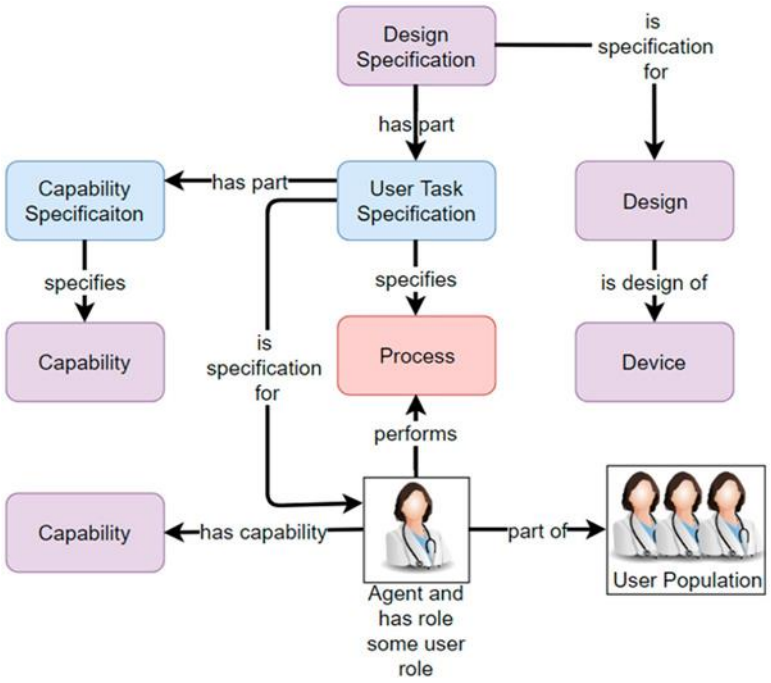

Figure 37. Representing user capabilities (source: Hagedorn et al. 2019).

This directs our attention to the realm of design requirements and design specifications; i.e., informational artefacts that specify how device designs are aligned (or perhaps misaligned) with the capabilities of particular user communities.

A recent example of how to represent capabilities from a design-oriented perspective is provided by Hagedorn et al. (2019). While Hagedorn et al. (2019) are primarily concerned with the design of medical instruments, their overall approach is one that can be applied to situations involving the design of IoT devices, including situations where the design of a device is informed by security-related objectives (as opposed to, for example, usability objectives).

The key features of Hagedorn et al.'s (2019) approach are depicted in Figure 37. In this diagram, the elements labelled design specification, user task specification, and capability specification are classes that prescribe some state-of-affairs. For this reason, they draw on the various classes that are situated under the heading of information content entity (see Section 3.6). In essence, specifications are information content entities that specify what could be the case or what ought to be the case. A user task specification, for example, specifies that a particular sort of process will be performed by a particular sort of (human) agent, and this process depends on the presence of a capability that the agent is deemed to possess. The capability, in this case, is represented by a capability specification, which is a description of what capabilities the agent is deemed to possess or the capabilities that they ought to acquire (perhaps via their participation in a training process).

Thus far, we have discussed how human factors might be represented in a BFO-conformant ontology. But this discussion is pitched at a level of abstraction that ignores many of the peculiarities associated with the domain of IoT security. The analytic problem, recall, was not just a problem of understanding what a human factor is; it is also a problem of understanding what human factors are relevant to the ontological characterization of IoT devices in a security-related context. The question, then, is what qualities, capabilities, and other entities ought to feature as part of an ontology of IoT security?

In response to this question, it is worth looking at the entities covered in existing human factors ontologies. Of particular interest is the Human Factors Ontology (HUFO) ontology described by Oltramari et al. (2015). What makes this ontology interesting is that it was designed to represent human factors in a cyber-security context. Given this, the contents of the ontology are poised to inform our understanding of the sorts of entities that might be relevant to an ontology of IoT security.

The key components of HUFO are as follows:

- **Social-Cognitive Characteristics:** Characteristics pertaining to an individual's personality, ideology, and ethical disposition.
- **Behavioural Characteristics:** Characteristics related to behaviour. Examples are said to include the likes of integrity, motivation, and rationality.
- **Knowledge/Skill Characteristics:** This includes things like expertise, proficiency, and comprehension.
- **Mental Health Characteristics:** Examples include things like mental stress, mental acuity, and emotional state.
- **Physical Health Characteristics:** Examples include physiological stress, age, and amount of sleep.
- **External/Environmental Characteristics:** This includes things like workload and authorised/unauthorised access.

From an ontological perspective, it has to be said that this is a somewhat mixed bag of entities. It seems reasonable to assume that the term "characteristic" is referring to something like a property or feature or attribute of something. In this sense, all the characteristics in the aforementioned list would (or perhaps should) fall under the general metaphysical heading of specifically dependent continuants. Some of these characteristics are relatively easy to situate within the ontological framework of BFO. Skills, for example, are understood to be capabilities (hence dispositions) that are realized in particular processes. Other characteristics are somewhat harder to classify. The term "amount of sleep," for example, is suggestive of a particular quantity or amount of something. That would put it in the class of qualities. The problem is that qualities are specifically dependent continuants that inhere in independent continuants. This suggests that sleep must be an independent continuant, but that is counterintuitive. In particular, it doesn't seem appropriate to regard sleep as a physical thing, like a chair, a table, or an iPhone. Rather than being an independent continuant, sleep is a process—something we do or something we participate in. That puts it in the metaphysical category of occurrents, which are logically disjoint from independent continuants. Presumably, what is implied by the term "amount of sleep" is the idea that an individual has participated in a sleeping process for a certain period of time, and the temporal extent of this period (the duration) represents the amount of sleep that an individual has had. At this point, we are into the realm of occurrent entities, such as processes and temporal regions, rather than dependent and independent continuants. Insofar as continuants feature as part of this representational picture, then the relevant entities are those situated under the heading of information content entities. Specifically, we are talking about the information derived from some sort of observation or measurement of the sleeping process. The value associated with his information content entity is the thing that tells us how 'much' sleep an individual has had.

While human factors are the primary ontological target of Oltramari et al. (2015), they are also interested in issues of trust. In fact, trust plays an important role in determining the scope of Oltramari et al.'s (2015) effort. The HUFO, they suggest, "illustrates the individual characteristics, situational characteristics, and relationships that influence the trust given to an individual." In view of this, it seems likely that Oltramari et al. (2015) are concerned with those features of a human individual that lead to that individual being regarded as trustworthy. In the wider literature, it is common to distinguish between trustors and trustees, with the former corresponding to the individual who places their trust in an individual, and the latter being the individual who is trusted. Schematically this can be represented as X (trustor) trusts Y (trustee) to ϕ, where ϕ denotes the thing that Y is trusted to do (or perhaps not to do). Oltramari et al.'s (2015) concern, it seems, is with the forces and factors that make Y seem trustworthy to X. That is to say, the overarching concern is with the features that lead X to believe that Y is trustworthy. As Oltramari et al. (2015) note, research suggests that there are a number of features that influence judgements of trustworthiness. Perhaps the most widely set of features are those described by Mayer et al. (1995), namely, ability, benevolence, and integrity.

While the formulation X trusts Y to ϕ captures the essence of trust relationships, it tends to overlook an important distinction between two kinds of trust (Reiersen 2017). The first of these is what might be called behavioural trust (or trust-as-action). This is the sort of trust that is play when X places their trust in Y, which is to say that X comes to rely on Y to ϕ and the performance (or non-performance) of ϕ has some discernible consequence (or impact) for X. In addition to behavioural trust, there is an attitudinal form of trust, where trust is conceptualized as (e.g.) a belief that X has about Y. This is the sort of trust that is typically discussed by trust theorists, especially those who operate from a social scientific or philosophical position (e.g., Hardin 2002). While there is a rather

obvious link between attitudinal and behavioural trust—in order for X to place their trust in Y, it seems reasonable to think that Y is deemed trustworthy by X—it is important to bear in mind that attitudinal and behavioural trust are not the same. From an ontological perspective, behavioural trust is an intentional or deliberate act, and this puts it in the metaphysical category of occurrents. By contrast, attitudinal trust is a property of the trustor—it is a belief that is held by a trustor about a trustee. This puts attitudinal trust in the realm of specifically dependent continuants, which are disjoint from occurrents.

To help reinforce the idea that attitudinal trust is distinct behavioural trust, it is worth considering that one can believe that someone is trustworthy without necessarily placing their trust in them. If X believes that Y is trustworthy, then X might be disposed to place their trust in Y, but they may have no need to do so, in which case the act of placing trust will not occur (see Hardin 2001, for more on this).

In addition to the distinction between attitudinal trust and behavioural trust, it is also worth bearing in mind that there is distinction between trust and trustworthiness. X may believe that Y is trustworthy, and they may place their trust in Y for this reason. Unfortunately, none of this means that Y is actually trustworthy. To be sure, if X *knows* that Y is trustworthy, then Y must qualify as trustworthy, for it is a basic condition of knowledge that one's beliefs should be aligned with the factive structure of reality (see Pritchard 2009). In general, however, the beliefs that X has about Y will often fall short of knowledge—they will count as nothing more than mere beliefs about the sort of entity that Y is. Whether those beliefs are true or false is an entirely different matter, and it is precisely this issue that lies at the heart of much of the academic debate about trust. The primary problem of trust is, in short, the problem of *knowing* whether a particular entity (the Y) qualifies as trustworthy.

At this point, it should be clear that the focus of Oltramari et al.'s (2015) effort is perhaps a little misplaced. To be sure, it is both interesting and important to understand the forces and factors that lead someone to believe that another entity is trustworthy. This, however, is not the same as understanding what it is that makes an entity deserving of our trust (i.e., trustworthy). One of the goals in security-related contexts is to understand the conditions under which a particular entity may be deemed to be trustworthy by another entity. But another, equally important, objective is to understand the conditions under which an entity *ought* to be trusted (or distrusted) on the grounds that they are actually trustworthy (or untrustworthy). These two objectives, it should be clear, are not the same. Given a good understanding of what leads someone to believe that another entity is trustworthy, a malignant actor could leverage this understanding to manipulate someone into believing they are trustworthy when, in fact, they are not trustworthy. The mere fact that the unfortunate victim believed they were interacting with someone trustworthy is neither here nor there as regards the actual presence of trustworthiness.

It seems, then, that an ontology of human factors will need to do more than list the forces and factors that shape an individual's trust-related beliefs and actions in respect of a particular trustee. In addition to this, we need to consider what it means for some entity to count as either trustworthy or untrustworthy. This looks to be important in a security context, for it is difficult to ascertain the hazards of X trusting Y to ϕ in the absence of an understanding of whether or not Y is trustworthy. If ϕ is a process that could have serious implications for X, then the trustworthiness of Y is a key concern (independent of X's beliefs about Y). If Y is trustworthy, then the risks to X are attenuated, for Y will at least attempt to do what they are being trusted to do. This will not be the case if Y should

prove to be untrustworthy. In this case, the act of placing trust is the trigger for a cascade of events that are unlikely to be in X's interest.

In order to provide an ontological characterization of trustworthiness, we first need to understand what trustworthiness is (or, at least, what it might be). While the nature of trustworthiness remains a topic of ongoing debate and discussion, there are at least some general points of agreement. One point of agreement is that trustworthiness is a property of things. In particular, it is property of the thing that is trusted, which is to say it is a property of the trustee. In fact, we can go beyond this and assert that trustworthiness is a dispositional property. This is, at least, the way that trustworthiness is typically conceptualized in the prevailing literature (Carter, in press; O'Hara 2012). Given this, we can assert that trustworthiness—whatever else it might—belongs to the metaphysical category of realizable entities. It is, in particular, a type of disposition that inheres in an independent continuant (e.g., a person, an organization, or a technological system).

By itself, the status of trustworthiness as a dispositional property does not tell us much about what it means to be trustworthy. We can, however, draw on the prevailing literature to advance our understanding of what it could be. The first thing to say is that dispositional properties are typically distinguished from categorical properties. In BFO, dispositional properties are represented by the **`DISPOSITION`** class and categorical properties are represented by the **`QUALITY`** class. While categorical properties are typically things that we can observe in the here-and-now, a dispositional property refers to things that could happen or that might happen given certain conditions. There is a large and sprawling literature regarding dispositional properties, which we won't have space to discuss here. Some notable introductions to the literature include Mumford (1998) and McKitrick (2018). For present purposes, we can simply acknowledge the idea that categorical properties are features or attributes or characteristics that can be observed or measured in the here-and-now (categorical properties are always manifest). Dispositional properties, by contrast, are not things that are easily measured in the absence of some sort of realization of the disposition. My ability to write computer code, for example, is a dispositional property, but it is not something that you can easily measure in the absence of some sort of test—where the notion of a test requires me to participate in a process that entails the manifestation (or realization) of the relevant ability.

At this point, we can draw on some terms from the realm of dispositional philosophy. This will help us better understand how to represent trust-related entities in an ontology. The most relevant terms are as follows (see McKitrick 2018):

- **Disposition Bearer:** An entity that possesses (bears) a dispositional property. In BFO, such entities are independent continuants.

- **Dispositional Property:** A class of properties that refer to potentialities. Dispositional properties include the likes of propensities, vulnerabilities, tendencies, capacities, capabilities, functions, and abilities. Some canonical examples are the fragility of a vase, or the solubility of salt in water. Dispositional properties inhere in objects that are the bearers of those dispositions. A fragile vase, for example, is an object (independent continuant) that bears the dispositional property of fragility.

- **Disposition Ascription:** The ascription of a dispositional property to a disposition bearer. This could be conceptualized as a process—i.e., the process of ascribing a disposition to a disposition bearer—or a belief about some entity's status as the bearer of a given disposition.

- **Disposition Manifestation:** The actualization, manifestation, realization, exercise, or execution or a dispositional property. This is an occurrent entity (e.g., a process).

- **Disposition Trigger:** An occurrent entity that (causally) triggers a disposition manifestation. The fragility (disposition) of the vase (independent continuant) is manifest in a breaking process (disposition manifestation) that is triggered by a dropping process (disposition trigger).

In respect of trustworthiness, we have already established the status of trustworthiness as a dispositional property. Dispositional properties inhere in disposition bearers, which, in the case of trustworthiness, will be the entity that is trusted or distrusted (i.e., the trustee). According to the above, a disposition ascription is characterized as either a process or a belief. This could either be the process that leads a trustor to believe that a trustee is trustworthy (or untrustworthy) or it could refer to the actual belief that is formed as a result of this process. In either case, the focus here is the trustor, not the trustee. That is to say, we are either talking about processes in which the trustor participates, or we are talking about the beliefs that are held by the trustor. These beliefs will refer to (or be about) the trustee (or disposition bearer). Specifically, the trustor's beliefs will refer to the trustworthiness of the trustee, or, equivalently, a dispositional property that inheres in a disposition bearer.

Disposition manifestations are the processes that reflect the realization of a dispositional property. For trustworthiness, the disposition manifestation will be a process that realizes the trustworthiness of the trustee. This will be a process in which the trustee (the disposition bearer) is a participant. For the most part, such processes will be those in which a trustee fulfils (or fails to fulfil) the trust that is placed in them. So, if X places their trust in Y, and Y is genuinely trustworthy, then Y will respond to the placement of trust by participating in a process that fulfils the trust that is placed in them. Assuming the earlier schematic of X trusts Y to ϕ, then Y 's trustworthiness is manifest in the fact that they perform ϕ, which is to say, they do what they are trusted to do, or they do what X expects them to do.

The only remaining term is disposition trigger. According to Ray et al. (2016), a disposition trigger is a process that is causally linked to the process in which a disposition is realized. For trustworthiness, the process in which the relevant disposition (trustworthiness) is realized is the process of fulfilling trust. The disposition trigger must therefore be a process that causes this process to occur. Thus, if process (P1) causes process (P2), and P2 is the realization of a disposition (D), then P1 is the disposition trigger for D (or, conversely, D is triggered by P1). If P2 is the process of fulfilling trust, then P1 must be the process that causes this process to occur. That process is, we suggest, best understood as the process corresponding to the placement of trust. This process is, of course, performed by the trustor. So, the general sequence of events is something like the following:

- X believes that Y is trustworthy. [In BFO, the belief will be represented as a mental quality and thus a quality (see Limbaugh et al. 2020).][32]

[32] Note that X's belief that Y is trustworthy may entail that X is disposed to place their trust in Y. The realization of this disposition would be the act of X their placing trust in Y. This sort of disposition is what might be called a trust disposition (see Mayer, Davis, and Schoorman 1995).

- X places their trust in Y. More specifically, X trusts Y to ϕ. [In BFO, the act of placing trust is a process.]
- The act of placing trust triggers the instantiation of a process (i.e., ϕ) in which Y is a participant. [In BFO, the act of placing trust is the disposition trigger.]
- In performing ϕ, Y manifests their trustworthiness and thereby fulfils the trust that was placed in them. [In BFO, ϕ is a process in which Y is a participant.]

The foregoing gives us a rudimentary way of representing trust-related entities in BFO. At the very least, we know where various trust-related terms ought to be situated within the BFO hierarchy. What is more, it is now clear where the work by Oltramari et al. (2015) fits within this emerging ontological framework. Oltramari et al., recall, are concerned with the factors that motivate beliefs about another's trustworthiness. These factors can be understood as the trustor's beliefs about the properties of a trustee, as well as the properties of the situation in which the act of placing trust might be performed (or the trustworthiness of the trustee realized). Note that these are beliefs held by the trustor; they are not the actual properties of the trustee (or the situation). The beliefs might be about these properties, but they are not the properties themselves. Such beliefs participate in processes that lead a trustor to form another belief about the trustworthiness of the trustee. In BFO, we would say that the trustor is a participant in a belief-forming process that involves other beliefs (as participants). The output of this process is a further belief whose content refers to the (perceived) trustworthiness of the trustee.

What about the factors that underwrite the trustee's status as a trustworthy entity? That is to say, what sort of things are indicative of the fact that the trustee is the bearer of a specific dispositional property called trustworthiness? We have explored this issue at length in earlier work, including that undertaken in respect of the CP-SOCIAM and PETRAS-DSF projects (Smart, Hall, and Boniface 2022b, 2022a). Suffice to say, there are a multiplicity of factors to consider here, many of which relate to the possession of dispositional properties other than trustworthiness, such as abilities. This establishes a point of contact with the present work, which we suggest could be taken forward as a future work activity (see Section 7.5).

Before departing the realm of dispositional properties, it is worth bearing in mind that multiple kinds of dispositional property have been discussed in the philosophical literature. We have already discussed the status of trustworthiness and capabilities as dispositional properties. Other dispositional properties include the likes of abilities, capacities, propensities, tendencies, and so on (see Mumford 1998). Arp and Smith (2008) mark a distinction between dispositions and tendencies. They suggest that:

> [A tendency] is a realizable dependent continuant that potentially (not invariably or definitely) causes a specific process in the object in which it inheres when the object is introduced into certain specific circumstances as a result of the object's physical structure property. […] A patient may have a tendency, and not a disposition, to commit suicide; while a crystal vase has a disposition, and not a tendency, to break when it hits the ground after being dropped from a tall building. We are referring to tendencies when we refer to genetic and other risk factors for specific diseases. (Arp and Smith 2008, p. 3)

The core distinction between dispositions and tendencies thus appears to relate to the 'sure-fire' nature of dispositions: a disposition is something that is invariably realized in the right conditions, while a tendency is a realizable entity whose realization is not guaranteed (see Jansen 2007, for more on tendencies). This sort of distinction is apt to be relevant in a security context, especially where we are talking about an individual's tendency to do certain things, such as engage in malicious activity. Admittedly, the distinction between tendencies and dispositions is not particularly clear-cut; nevertheless, there may be some value in attending to this distinction in future work. At first sight, a cyber-criminal may seem to be an individual who is disposed to participate in criminal activity given the presence of certain background conditions. On other hand, it may be better to characterize the individual as one who is the bearer of a cyber-criminal tendency, as opposed to a disposition. Cyber-criminals are, we may suppose, those individuals who are inclined to exploit an opportunity should one become available, but this does not mean they are inclined to exploit every such opportunity. Insofar as we adopt the distinction proposed by Arp and Smith (2008), then this would count as a tendency towards cyber-criminal behaviour in certain (criminogenic) environments (see Ward and Stewart 2003), not a disposition towards such behaviour.

7 RECOMMENDATIONS

In this section, we outline our recommendations in respect of future work. These recommendations are informed by the results of our survey of existing ontologies, as well as attempts to provide an ontological characterization of common security concepts.

7.1 Ontology Design Principles

Our first recommendation pertains to the design of security ontologies. We suggest that attempts to develop security ontologies should adhere to a common set of design principles that are intended to promote reuse, interoperability, comprehension, and critical evaluation. Some insight into these principles is provided by the following list, which is owed to Babcock et al. (2021):[33]

1. Ontologies should use a well-specified syntax and share a common space of identifiers.
2. Ontologies should be openly available in the public domain for reuse.
3. Ontologies in neighbouring domains should be developed in a collaborative effort.
4. Ontologies should be developed in a modular fashion.
5. Ontologies should have a clearly specified scope.
6. Ontologies should use common unambiguously defined relations between their terms.
7. Ontologies should conform to a common top-level architecture.

Conformance to a common top-level architecture is particularly important. We recommend that all security ontologies should be developed as extensions to an upper- or top-level ontology. In the present report, we have adopted BFO as the top-level architecture for our modelling efforts. There are, however, other top-level ontologies that could be used, such as UFO, DOLCE, and GFO (see Partridge et al. 2020, for a recent survey of upper ontologies).

7.2 Modularity

We recommend that security ontologies be designed in a modular fashion with each ontology focusing on a restricted range of terms or concepts. This recommendation is intended to simplify ontology development efforts, promote interoperability and support cross-disciplinary integration.

A nice example of modularity stems from the attempt to develop an ontology of COVID-19. Figure 38 shows how a COVID-19 ontology (in this case, IDO-COVID-19) is built on top of other ontologies, each of which describes a particular part of reality at a different level of abstraction (see Babcock et al. 2021). The IDO-COVID-19 ontology is thus developed as an extension of the Coronavirus Infectious Disease Ontology (CIDO), which is, itself, built on top of the Virus Infectious Disease Ontology (VIDO). The virtues of this approach should be relatively clear. By relying on an existing ontology, the ontology development effort is greatly simplified. Consider, for example, that one could develop an ontology for other viral diseases simply by extending the VIDO ontology and adhering to the modelling conventions adopted in both the CIDO and IDO-COVID-19 ontologies.

[33] For more on BFO best practice principles, see Arp et al. (2015, chap. 4).

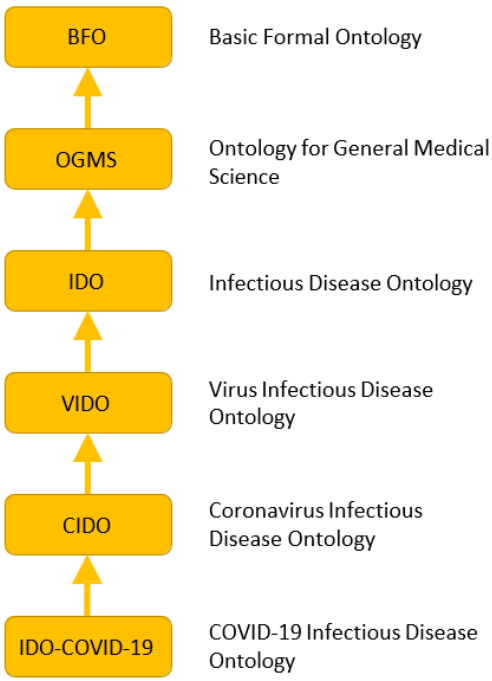

Figure 38. The vertical integration of ontologies for COVID-19. [Arrows symbolize extension relations, such that OGMS extends BFO.]

Figure 38 depicts what we will call vertical integration—a state-of-affairs in which one ontology relies on a suite of progressively more abstract ontologies. This highlights one of the virtues of a commitment to modularity. There is, however, a second form of integration that also benefits from a commitment to modularity. This is what we will call horizontal integration. Horizontal integration occurs when one ontology relies on ontologies that were developed for, perhaps, entirely different domains, domains that may, at first sight, seem utterly unrelated to the particular domain in which one is interested. Consider, for example, that there are both terminological and conceptual correspondences between the security and medical domain. One such correspondence arises in respect of the notion of risk. Just as there are security risks, so too there are health risks (e.g., the risk of developing a particular disease if one should engage in a certain form of behaviour). Another form of correspondence arises in respect of security mechanisms and health interventions. Just as a security mechanism seeks to reduce the risk associated with an undesirable outcome, so a health intervention seeks to reduce the risk of developing a particular disease (also an undesirable outcome). One way to minimize one's risk of developing an infectious disease, for example, is to receive a vaccine that prevents one from participating in a disease course that represents the realization of a disease. Similarly, the introduction of a security mechanism might be seen as a way of preventing or blocking a cyber-attacker from participating in certain processes.[34] There are, of course, other forms of correspondence here. The notion of a virus, for example, is typically

[34] The notion of an entity being blocked from participating in processes occurs in both the medical and security domains (Goldfain, Smith, and Cowell 2010; Oliveira et al. 2022). In particular, Goldain et al. (2010) rely on the notion of a blocking disposition as a means of advancing our ontological understanding of disease resistance. A similar appeal to blocking dispositions can be found in the security literature (e.g., Oliveira et al. 2022). For a more general approach to modelling prevention, see Baratella et al. (2022).

understood with regard to the realm of biological entities (biological viruses). But there are also computer viruses, and the hazards of computationally-inflected forms of contagion are no less serious for society than is the threat of a virally-mediated pandemic.

The point here is that there are certain concepts that appear to be relevant to multiple, ostensibly disparate, domains. The notions of threat, risk, harm, vulnerability, impact, and so on are, as we noted in Section 4.1, recurring terms that appear across multiple security ontologies. In fact, however, these terms are not confined to the realms of cybersecurity, nor are they the sole preserve of those with an interest in cybersecurity. Instead, these terms are ones that recur across *multiple* disciplines. Given this, it is not particularly clear why abstract terms like threat, risk, vulnerability, would even appear in a domain-specific security ontology, for these terms really ought to be defined in (and curated by) ontologies (and ontology authors) whose interests cross-cut multiple areas of research. In short, what is required is an ontology of risk that focuses *solely on risk*, not on the way that risk is understood in different disciplinary contexts. Given the availability of this ontology, it will then be possible to develop more specialized ontologies that specify what it means for someone to be exposed to a risk factor (in medicine) or what risk means in a security-related context. The effort to develop these more specialized risk ontologies is greatly simplified by the presence of a generic risk ontology, just as the development of a COVID-19 ontology is greatly simplified by the presence of a more abstract ontology of virally-transmitted infectious diseases. In fact, in the majority of cases, we would not expect the more specialized ontologies to be particularly concerned with the semantics of the term risk, for such semantics should be resolved really be resolved in a manner that papers over the more specialized (domain-specific) uses of the risk concept.

7.3 Modelling Patterns

As a means of supporting standardized approaches to data modelling, we advocate the development of a library of modelling patterns, similar to those that have been presented throughout this report. The aim here is to demonstrate *how* common scenarios and states-of-affairs ought to be represented from an ontological standpoint. The provision of concrete examples is, we suggest, a useful accompaniment to the more abstract characterizations of terms and concepts that are presented in the ontological literature. It is, in particular, important to understand how an ontology can be used for the practical project of representing common bodies of data. Consider, for example, the situation where one wants to represent the reliability of an IoT sensor. While the description of ontology classes (in CCO or elsewhere) might provide some insight into how to resolve this problem, our own experience suggests that these descriptions are seldom sufficient for important modelling decisions. [And this assumes that such descriptions are even available. In the majority of cases, they aren't.] At the very least, the human-readable descriptions leave too many questions unanswered; the description of ontology elements is often (perhaps unsurprisingly) directed to the elucidation of these ontology elements. But the elucidation of an ontology element is not the same as an understanding of how that ontology element ought to be used as part of practical data modelling efforts. One could have a robust conceptual understanding of what makes something a member of the class of IoT sensors, and one could also have a robust understanding of what it means for something to be reliable. But such forms of understanding do not tell us how to represent the specific reliability of a specific IoT sensor at different points in time. That sort of understanding is, we suggest, best supported via the provision of concrete examples.

We make no claims as to the precise format in which a library of modelling patterns is to be delivered. We do, however, recommend that each modelling pattern should be accompanied by UML

(object) diagrams, a textual narrative, sample SPARQL queries, and (ideally) an OWL-based serialization of the objects featured in the modelling pattern.

7.4 Ontology Repositories

Unlike the medical domain, there are no common repositories for security ontologies. This limits the extent to which security ontologies can be shared, accessed, queried, and critiqued. In future work, we recommend the creation of a global repository for Internet-related ontologies, with IoT security ontologies featuring as one of the areas covered by the repository. Some insight into the nature of this proposed repository can be gleaned from a consideration of the following biomedical repositories:

- NCBO BioPortal (https://bioportal.bioontology.org/)
- Ontology Lookup Service (https://www.ebi.ac.uk/ols/index)
- Ontobee (https://ontobee.org/)

Each of these repositories enables interested partners to browse, search, and download ontologies. What is more, many of the ontologies hosted by these repositories conform to a particular top-level ontology, namely, BFO. At a minimum, we suggest that an ontology repository for the Internet (which subsumes the IoT) should provide similar capabilities. We also recommend that the repository provide support for community discussion and collaboration with ontology authors. Each ontology hosted by the portal must be available for download. The convention in bio-medicine is to make ontologies available via GitHub in addition to including them in the repository.

7.5 Trust, Trustworthiness, and Human Factors

In our view, the ontological characterization of IoT devices, CPSs, digital twins, is not particularly problematic. While there might be disagreements about the semantic characterization of these things, their membership of one or other sort of ontological category is not in any doubt. Whatever else an IoT device might be, it is, according to BFO, a type of material entity and thus a type of independent continuant. In this respect, the development of an IoT device ontology ought to be relatively straightforward.

What is more problematic are ontologies that deal with more abstract concepts, such as trust and trustworthiness. The problem here is that our conceptual understanding of these terms remains in its infancy, and it is thus unclear how these terms ought to be represented in an ontology. The same is true of any number of other concepts that are found in the human factors domain.

In Section 6, we provided the basis for the resolution of at least some of these problems. Our recommendation is that this work be taken forward via a concerted effort to develop (BFO-conformant) ontologies that explicate the current state-of-the-art when it comes to our understanding of concepts, such as knowledge, trust, trustworthiness, privacy, robustness, resilience, value, and objective. This is not something that can be tackled by a single discipline; instead, it requires a concerted effort that spans multiple disciplines. Crucially, it is vital that future work should incorporate the insights and ideas emanating from the realms of philosophy, for many of the aforementioned concepts have been the subject of sustained analytic scrutiny for many decades, and the conceptual work in this area tends to be more advanced than the current state-of-the-art in engineering disciplines.

7.6 Philosophical Engineering

Throughout the present report, we have encountered multiple appeals to the philosophical literature. In one sense, this is unsurprising, for the development of formal ontologies is really just an exercise in applied philosophy. While the bulk of domain-level ontologies tend to be developed by engineers, the majority of upper-level ontologies are informed by work in philosophy, specifically, work in metaphysics. To help us appreciate the importance of this link, it is worth bearing in mind that many of the classes we have discussed in previous sections are derived from work in philosophy. Examples include the likes of occurrents, continuants, dispositions and categorical properties (qualities). The same is true of many of the relations that form the backbone of ontology development efforts. Such relations include:

- **Dependence.** As we discussed in Section 3.2, BFO relies on the notion of existential dependence to distinguish independent continuants from dependent continuants.

- **Inherence.** Specifically dependent continuants (e.g., dispositions) are said to inhere in their bearer, which, in the case of BFO, are independent continuants.

- **Realization.** In BFO, realization assumes the form of a relation between a realizable entity and an occurrent entity (Arp and Smith 2008). In the wider philosophical literature, different kinds of realization have been identified. Examples include material realization, mechanistic realization, aggregate realization, and emergent realization (Wilson and Craver 2007).

- **Parthood.** Parthood is a specialist area of philosophical research. BFO includes relations to represent different kinds of parthood, including processual parthood and conventional mereology. There are, however, important distinctions between a number of mereological constructs. In philosophy, for example, a distinction is sometimes drawn between constitutive and compositional relations (e.g., Glennan 2021).

- **Grounding.** Dispositional properties are often understood to be grounded in categorical properties. This grounding relation is not explicitly represented in BFO; nevertheless, an understanding of grounding relations is often key to understanding the distinction between different types of realizable entity, most notably the distinction between dispositions and roles (see Arp and Smith 2008).

- **Participation.** Participation represents the primary relation between continuants and occurrents (see Rodrigues and Abel 2019). In BFO, continuants are said to participate in occurrents. Philosophers tend to distinguish between different kinds of participatory relationship, although such distinctions are seldom used in BFO ontologies (see Smith and Grenon 2004).

The proper use of these relations poses a challenge for those who are not conversant with the philosophical literature, and for this reason we suggest that future work should strive to solicit the support of the philosophical community, especially when it comes to the development of ontologies that extend an upper ontology.

Philosophical input is also apt to be valuable when it comes to our basic understanding of key terms and concepts. Trust and trustworthiness, for example, are two concepts that have been the subject of sustained philosophical scrutiny for several decades. Unfortunately, very little of this work is

reflected in ontologies that are intended to tackle the notion of trust. The result is a failure to learn from past mistakes and persist with ontological formulations that (from a philosophical standpoint) are known to be unworkable.

At the same time, there is plenty of work in engineering that fails to find its way into philosophical discourse, leading to debates and discussions that seem a little out of kilter with practical concerns. Philosophers, for example, tend to be preoccupied with human-based forms of trust, and they seldom venture into terrains where the primary concern is with technologically-inflected forms of trust. Our sense is that genuine progress in this (and other areas) requires a concerted effort to establish bi-directional forms of communication and influence that are apt to benefit both philosophical and engineering disciplines in about equal measure.

In addition to work on trust/trustworthiness, there is also a need to establish interdisciplinary connections in respect of value (see Section 4.2), risk (see Section 4.4), vulnerability (see Section 4.7), and security mechanisms (see Section 4.8). It will also be important, in future work, to determine how the practical effort to construct applied ontologies reshapes the trajectory of philosophical thought. Some insight into the nature of these challenges stems from the earlier discussion of vulnerabilities, where we explored the status of vulnerabilities as either internally- or externally-grounded dispositional properties. Crucially, this is one area where BFO seems to find itself caught in a dilemma, for BFO is committed to the status of vulnerability as an internally-grounded disposition, and this does not appear to be consistent with our intuitions about the relativistic status of vulnerabilities (i.e., the fact that vulnerabilities vary according to situational factors, as opposed to shifts in intrinsic properties).

7.7 Active Inference

Our final recommendation builds on the aforementioned appeal to interdisciplinary cooperation by suggesting that greater attention ought to be devoted to the emerging theoretical framework of active inference and free energy minimization (Parr, Pezzulo, and Friston 2022). Work in this area actually goes by a variety of names, such as predictive coding (Rao and Ballard 1999), predictive processing (Clark 2013, 2016), the free energy principle (Friston 2009, 2010), active inference (Friston et al. 2015; Friston et al. 2017; Parr, Pezzulo, and Friston 2022) and the Bayesian brain (Doya et al. 2011). In all cases, however, the general idea is that biological brains can be understood as hierarchically-organized (multi-layered) networks that are constantly striving to minimize prediction error. As it turns out, this simple imperative to minimize prediction error provides the basis for a unified explanatory account of multiple psychological phenomena. It also serves as a common organizational principle for multiple types of machine learning systems, especially those that avail themselves of generative models. In this sense, then, the active inference framework provides a natural point of contact between work in both generative AI and cognitive neuroscience. It also provides a means of anchoring a host of seemingly mysterious concepts to a common physical bedrock. From an active inference perspective, for example, trustworthiness is just a disposition to minimize a socially-inflected variant of (neurally-realized) prediction or neural free energy. It is, in particular, a disposition to minimize the error associated with a trustor's expectations about the future. The trustworthy are those who work to minimize this error, either by doing what the trustor expects them to do or by modifying the trustor's expectation so as to bring it into alignment with the fact structure of reality (Smart, Hall, and Boniface 2022a). The point here is that we can understand (and model!) the dynamics of trust in pretty much the same way we understand (and model) any of the other phenomena that are tackled by the active inference framework.

But it isn't just trust and trustworthiness where the active inference framework comes in useful. According to active inference, biological brains have evolved to minimize the quantity of information that is traded by anatomically connected neural regions. Such forms of optimization are not particularly surprising, given that biological brains are subject to a set of energetic and computational constraints that are at least no less severe than those faced by a conventional IoT device. In this respect, closer attention to neuromorphic (or neuromimetic) forms of information flow could yield practical solutions for dealing with resource-constrained IoT devices.

Active inference also yields the potential for a revolutionary shift in our approach to the development of CPS and digital twins. As Jarwar et al. (2022) remark, IoT devices "are the ears and eyes for Cyber Physical Systems (CPS) and smart applications." The reference to eyes and ears is no doubt intended to be analogical here, but we would like to suggest that there is a more literal way of understanding these biological references. In particular, we suggest that we can understand the relationship between IoT devices, CPSs, and digital twins in more or less the same way that we understand the dynamics of information processing and (crucially) agential control in the active inference framework. From an active inference perspective, an IoT sensor is a source of sensory inputs, and the computational prerogative of a CPS is to predict the statistical structure of the inputs (the data) that emanates from this sensor. This predictive capability is engendered by a generative model that embodies the causal structure of the particular part of sensory reality (the sensorium) in which the CPS is embedded. From a mechanistic standpoint, this generative model features as one of the components of a digital twin whose goal is pretty much the same as that of the biological brain. That is to say, the goal of a digital twin is to represent reality in such a way that enables it to predict the incoming stream of sensory data. But this is not the only goal of a digital twin, for there are two ways one can increase the accuracy of one's predictions about the future. Firstly, one can focus on improving the quality of one's predictions by learning from one's past mistakes (or prediction errors). Such learning involves changes in the structure of one's generative model, such that the model is better placed to issue more accurate predictions in the future. But there is a second way one can minimize prediction error. Biologically one can implement actions so as to ensure that the future evolves in a manner that is consistent with one's predictions. If one predicts that one will be at the supermarket at such-and-such a time, then one can fulfil this prediction by (e.g.) driving to the supermarket at that time. By doing so, one has, in effect, engineered a self-fulfilling prophecy—one has used one's own bodily resources to bring about or create or generate one's own sensory future. Much the same, we suggest, applies to the relationship between digital twins, CPS, and IoT actuators. A digital twin is like the biological brain, in the sense that it contains a generative model that predicts the incoming sensory signal. A CPS is like the body of an individual. It is an interconnected system of sensors and actuators that work together so as to deliver sensory flows to the brain (perception) and (via action) manipulate those flows in such a way as to conform to the brain's predictions. The active manipulation of these sensory flows, in an IoT context is accomplished by IoT actuators.

The end result is an integrated vision of CPSs that draws on our current best understanding of both the biological brain and generative AI systems. Fleshing out this vision will no doubt require the resolution of multiple problems; but it is, we suggest, an important and compelling avenue for future research. Ultimately, the vision is one in which the next generation of CPSs will function as the intelligent controllers of all manner of social, technological, and environmental processes. And their overarching prerogative will be the same as that of any other intelligent system—to coordinate one's present actions with a set of future possibilities, such that one's actions in the here-and-now are the stepping stones to a future that one 'expects' oneself to be in. As human individuals, our brains optimistically 'expect' themselves to be in a certain future—one in which we are warm, well-fed, and

(perhaps most important of all) socially-connected (see Baumeister and Leary 1995; Beckes and Coan 2011; Beckes and Sbara 2022). But what sort of expectations will be harboured by the next generation of CPSs? What futures will their actions forge? And how can we secure the future of those who will shortly come to inherit our socio-technical legacy? In all likelihood, such questions cannot be addressed via the efforts of a single academic discipline. As with the project of developing ontologies to represent security-related concepts, the project of engineering the next generation of CPSs is one that is likely to require a coordinated effort that transcends multiple disciplines. Futurology is always difficult, as is inter-disciplinary collaboration; nevertheless, it is arguably important that we confront such difficulties if we, or rather our descendants, are to inhabit a future that is worth expecting.

REFERENCES

Alanen, Jarmo, Joonas Linnosmaa, Timo Malm, Nikolaos Papakonstantinou, Toni Ahonen, Eetu Heikkilä, and Risto Tiusanen. 2022. "Hybrid ontology for safety, security, and dependability risk assessments and Security Threat Analysis (STA) method for industrial control systems." *Reliability Engineering & System Safety* 220:108270.

Albarracin, Mahault, Inês Hipólito, Safae Essafi Tremblay, Jason G Fox, Gabriel René, Karl Friston, and Maxwell JD Ramstead. 2023. "Designing explainable artificial intelligence with active inference: A framework for transparent introspection and decision-making." *arXiv Preprint* arXiv:2306.04025, https://doi.org/10.48550/arXiv.2306.04025.

Arp, Robert, and Barry Smith. 2008. "Function, role, and disposition in Basic Formal Ontology." *Nature Precedings*, 1–4.

Arp, Robert, Barry Smith, and Andrew D Spear. 2015. *Building Ontologies with Basic Formal Ontology*. Cambridge, Massachusetts, USA: MIT Press.

Babcock, Shane, John Beverley, Lindsay G Cowell, and Barry Smith. 2021. "The infectious disease ontology in the age of COVID-19." *Journal of Biomedical Semantics* 12 (Article 13): 1–20.

Baratella, Riccardo, Mattia Fumagalli, Ítalo Oliveira, and Giancarlo Guizzardi. 2022. "Understanding and modeling prevention." In *International Conference on Research Challenges in Information Science*, edited by Renata Guizzardi, Jolita Ralyté, and Xavier Franch, 446:389–405. Barcelona, Spain: Springer.

Baumeister, Roy F. and Leary, Mark R. 1995. "The Need to Belong: Desire for Interpersonal Attachments as a Fundamental Human Motivation." *Psychological Bulletin* 117 (3): 497-529.

Beckes, Lane and Coan, James A. 2011. "Social baseline theory: The role of social proximity in emotion and economy of action." *Social and Personality Psychology Compass* 5 (12): 976–988.

Beckes, Lane and Sbarra, David A. 2022. "Social baseline theory: State of the science and new directions." *Current Opinion in Psychology* 43: 36–41.

Boholm, Åsa, and Hervé Corvellec. 2011. "A relational theory of risk." *Journal of Risk Research* 14 (2): 175–190.

Borgo, Stefano, Roberta Ferrario, Aldo Gangemi, Nicola Guarino, Claudio Masolo, Daniele Porello, Emilio M Sanfilippo, and Laure Vieu. 2022. "DOLCE: A descriptive ontology for linguistic and cognitive engineering." *Applied Ontology* 17 (1): 45–69.

Borgo, Stefano, Antony Galton, and Oliver Kutz. 2022. "Foundational ontologies in action." *Applied Ontology* 17 (1): 1–16.

Carter, J Adam. in press. "Trust and Trustworthiness." *Philosophy and Phenomenological Research*, 1–18.

Casola, Valentina, Rosario Catelli, and Alessandra De Benedictis. 2019. "A first step towards an ISO-based information security domain ontology." In *28th International Conference on Enabling*

Technologies: Infrastructure for Collaborative Enterprises (WETICE), 334–339. Capri, Italy: IEEE.

Ceusters, Werner, and Barry Smith. 2010. "Foundations for a realist ontology of mental disease." *Journal of Biomedical Semantics* 1 (Article 10): 1–23.

Chemero, Anthony. 2009. *Radical Embodied Cognitive Science*. Cambridge, Massachusetts, USA: MIT Press.

Cheong, Hyunmin, and Adrian Butscher. 2019. "Physics-based simulation ontology: an ontology to support modelling and reuse of data for physics-based simulation." *Journal of Engineering Design* 30 (10–12): 655–687.

Clark, Andy. 2006. "Material symbols." *Philosophical Psychology* 19 (3): 291–307.

Clark, Andy. 2013. "Whatever Next? Predictive Brains, Situated Agents, and the Future of Cognitive Science." *Behavioral and Brain Sciences* 36 (3): 181–204.

Clark, Andy. 2016. *Surfing Uncertainty: Prediction, Action and the Embodied Mind*. New York, New York, USA: Oxford University Press.

Clark, Andy. 2020. "Beyond desire? Agency, choice, and the predictive mind." *Australasian Journal of Philosophy* 98 (1): 1–15.

Compton, Michael, Payam Barnaghi, Luis Bermudez, Raul Garcia-Castro, Oscar Corcho, Simon Cox, John Graybeal, Manfred Hauswirth, Cory Henson, and Arthur Herzog. 2012. "The SSN ontology of the W3C semantic sensor network incubator group." *Journal of Web Semantics* 17:25–32.

Craver, Carl. 2007. *Explaining the Brain: Mechanisms and the Mosaic Unity of Neuroscience*. Oxford, UK: Clarendon Press.

Craver, Carl, and James Tabery. 2016. "Mechanisms in science." In *The Stanford Encyclopedia of Philosophy*, Spring 2016, edited by Edward N Zalta. Stanford, California, USA: Stanford University.

Craver, Carl F, and Lindley Darden. 2013. *In Search of Mechanisms: Discoveries Across the Life Sciences*. Chicago, Illinois, USA: The University of Chicago Press.

CUBRC. 2020a. *An Overview of the Common Core Ontologies*. Technical report. Buffalo, New York, USA: CUBRC, Inc.

CUBRC. 2020b. *Modeling Information with the Common Core Ontologies*. Technical report. Buffalo, New York, USA: CUBRC, Inc.

Da Costa, Lancelot, Pablo Lanillos, Noor Sajid, Karl Friston, and Shujhat Khan. 2022. "How active inference could help revolutionise robotics." *Entropy* 24 (Article 361): 1–7.

Daniel, Desiree. 2014. "Resilience as a Disposition." In *Formal Ontology in Information Systems*, edited by Pawel Garbacz and Oliver Kutz, 267:171–182. Frontiers in Artificial Intelligence and Applications. Rio de Janeiro, Brazil: IOS Press.

Donohue, Brian, Mark Jensen, Alexander P Cox, and Ron Rudnicki. 2018. "A common core-based cyber ontology in support of cross-domain situational awareness." In *Ground/Air Multisensor Interoperability, Integration, and Networking for Persistent ISR IX*, edited by Michael A Kolodny, Dietrich M Wiegmann, and Tien Pham, 10635:65–74. Orlando, Florida, USA: SPIE.

Doya, Kenji, Shin Ishii, Alexandre Pouget, and Rajesh P N Rao, eds. 2011. *Bayesian Brain: Probabilistic Approaches to Neural Coding*. Cambridge, Massachusetts, USA: MIT Press.

Drobnjakovic, Milos, Boonserm Kulvatunyou, Farhad Ameri, Chris Will, Barry Smith, and Albert Jones. 2022. "The Industrial Ontologies Foundry (IOF) Core Ontology." In *12th International Workshop on Formal Ontologies Meet Industry*, edited by Emilio M Sanfilippo, Mohamed-Hedi Karray, Dimitrios Kyritsis, and Arkopaul Sarkar, 3240:1–13. Tarbes, France: CEUR Workshop Proceedings.

Friston, Karl. 2009. "The free-energy principle: A rough guide to the brain?" *Trends in Cognitive Sciences* 13 (7): 293–301.

Friston, Karl. 2010. "The free-energy principle: A unified brain theory?" *Nature Reviews Neuroscience* 11 (2): 127–138.

Friston, Karl, Thomas FitzGerald, Francesco Rigoli, Philipp Schwartenbeck, and Giovanni Pezzulo. 2017. "Active inference: A process theory." *Neural Computation* 29 (1): 1–49.

Friston, Karl, Francesco Rigoli, Dimitri Ognibene, Christoph Mathys, Thomas Fitzgerald, and Giovanni Pezzulo. 2015. "Active inference and epistemic value." *Cognitive Neuroscience* 6 (4): 187–214.

Glennan, Stuart. 2017. *The New Mechanical Philosophy*. Oxford, UK: Oxford University Press.

Glennan, Stuart. 2021. "Corporeal composition." Synthese, 198 (12): 11439–11462.

Glennan, Stuart, and Phyllis McKay Illari, eds. 2018. *The Routledge Handbook of Mechanisms and Mechanical Philosophy*. New York, New York, USA: Routledge.

Goldfain, Albert, Barry Smith, and Lindsay G Cowell. 2010. "Dispositions and the infectious disease ontology." In *Formal Ontology in Information Systems*, edited by Antony Galton and Riichiro Mizoguchi, 400–413. Toronto, Canada: IOS Press.

Goldfain, Albert, Barry Smith, and Lindsay G Cowell. 2011. "Towards an ontological representation of resistance: the case of MRSA." *Journal of Biomedical Informatics* 44 (1): 35–41.

Guizzardi, Giancarlo, 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): 167–210.

Hagedorn, Thomas J, Barry Smith, Sundar Krishnamurty, and Ian Grosse. 2019. "Interoperability of disparate engineering domain ontologies using basic formal ontology." *Journal of Engineering Design* 30 (10–12): 625–654.

Haller, Armin, Krzysztof Janowicz, Simon JD Cox, Maxime Lefrançois, Kerry Taylor, Danh Le Phuoc, Joshua Lieberman, Raúlg García-Castro, Rob Atkinson, and Claus Stadler. 2019. "The modular

SSN ontology: A joint W3C and OGC standard specifying the semantics of sensors, observations, sampling, and actuation." *Semantic Web* 10 (1): 9–32.

Hardin, Russell. 2001. "Conceptions and Explanations of Trust." In *Trust in Society*, edited by Karen S Cook, 3–39. New York, New York, USA: Russell Sage Foundation.

Hardin, Russell. 2002. *Trust and Trustworthiness*. New York, New York, USA: Russell Sage Foundation.

Hastings, Janna, Werner Ceusters, Barry Smith, and Kevin Mulligan. 2011. "Dispositions and Processes in the Emotion Ontology." In *2nd International Conference on Biomedical Ontology*, edited by Olivier Bodenreider, Maryann E Martone, and Alan Ruttenberg, 71–78. Buffalo, New York, USA: CEUR-WS.org.

Illari, Phyllis McKay, and Jon Williamson. 2012. "What is a mechanism? Thinking about mechanisms across the sciences." *European Journal for Philosophy of Science* 2 (1): 119–135.

Jansen, Ludger. 2007. "Tendencies and other realizables in medical information sciences." *The Monist* 90 (4): 534–554.

Jarwar, Aslam, Jamie Tooth, and Jeremy Watson. 2022. *Secure ontologies for IoT Devices: A Review and Gap Analysis*. Technical report. London, UK: PETRAS National Centre of Excellence in IoT Systems Cybersecurity.

Jarwar, Muhammad Aslam, Jeremy Watson, Uchenna P Daniel Ani, and Stuart Chalmers. 2022. "Industrial Internet of Things security modelling using ontological methods." In *12th International Conference on the Internet of Things*, edited by Evangelos Niforatos, Gerd Kortuem, Nirvana Meratnia, Josh Siegel, and Florian Michahelles, 163–170. Delft, The Netherlands: ACM.

Jensen, Mark, Alexander P Cox, Brian Donohue, and Ron Rudnicki. 2018. "Problems with prescriptions: disentangling data about actual versus prescribed entities." In *Ground/Air Multisensor Interoperability, Integration, and Networking for Persistent ISR IX*, edited by Michael A Kolodny, Dietrich M Wiegmann, and Tien Pham, 10635:91–99. Orlando, Florida, USA: SPIE.

Kaiser, Marie I, and Beate Krickel. 2017. "The metaphysics of constitutive mechanistic phenomena." *The British Journal for the Philosophy of Science* 68 (3): 745–779.

Köhler, Sebastian, Sandra C Doelken, Christopher J Mungall, Sebastian Bauer, Helen V Firth, Isabelle Bailleul-Forestier, Graeme CM Black, Danielle L Brown, Michael Brudno, and Jennifer Campbell. 2014. "The Human Phenotype Ontology project: linking molecular biology and disease through phenotype data." *Nucleic Acids Research* 42 (D1): D966–D974.

Le, Duc-Hau, and Lan TM Dao. 2018. "Annotating diseases using human phenotype ontology improves prediction of disease-associated long non-coding RNAs." *Journal of Molecular Biology* 430 (15): 2219–2230.

Limbaugh, David, David Kasmier, Werner Ceusters, and Barry Smith. 2019. "Warranted Diagnosis." In *International Conference on Biomedical Ontology*, edited by Alexander D Diehl, William D Duncan, and Gloria Sanso, 2931:1–10. New York, New York, USA: CEUR Workshop Proceedings.

Limbaugh, David, Jobst Landgrebe, David Kasmier, Ronald Rudnicki, James Llinas, and Barry Smith. 2020. “Ontology and Cognitive Outcomes.” *Journal of Knowledge Structures and Systems* 1 (1): 3–22.

Loebe, Frank, Patryk Burek, and Heinrich Herre. 2022. “GFO: The General Formal Ontology.” *Applied Ontology* 17 (1): 71–106.

Machamer, Peter, Lindley Darden, and Carl F Craver. 2000. “Thinking about mechanisms.” *Philosophy of Science* 67 (1): 1–25.

Mandrick, Bill, and Barry Smith. 2022. “Philosophical foundations of intelligence collection and analysis: a defense of ontological realism.” *Intelligence and National Security* 37 (6): 809–819.

Mayer, Roger C, James H Davis, and F David Schoorman. 1995. “An integrative model of organizational trust.” *Academy of Management Review* 20 (3): 709–734.

McKitrick, Jennifer. 2018. *Dispositional Pluralism*. Oxford, UK: Oxford University Press.

Merrell, Eric, David Limbaugh, Peter Koch, and Barry Smith. 2022. Capabilities. https://philpapers.org/rec/MERC-14.

Mizoguchi, Riichiro, and Stefano Borgo. 2022. “YAMATO: Yet-another more advanced top-level ontology.” *Applied Ontology* 17 (1): 211–232.

Mozzaquatro, Bruno A, Ricardo Jardim-Goncalves, and Carlos Agostinho. 2015. “Towards a reference ontology for security in the Internet of Things.” In *2015 IEEE International Workshop on Measurements & Networking*, 1–6. Coimbra, Portugal: IEEE.

Mozzaquatro, Bruno Augusti, Carlos Agostinho, Diogo Goncalves, João Martins, and Ricardo Jardim-Goncalves. 2018. “An Ontology-Based Cybersecurity Framework for the Internet of Things.” *Sensors* 18 (9): 1–20.

Mumford, Stephen. 1998. *Dispositions*. Oxford, UK: Oxford University Press.

Naray, Rachel Kuzio de, Brian A Haugh, and Steven P Wartik. 2022. *An Ontology for the Embedded System TTP Matrix*. Technical report. Alexandria, Virginia, USA: Institute for Defense Analysis.

O’Hara, Kieron. 2012. *A General Definition of Trust*. Technical report. Southampton, UK: Electronics and Computer Science, University of Southampton.

Oliveira, Ítalo, Mattia Fumagalli, Tiago Prince Sales, and Giancarlo Guizzardi. 2021. “How FAIR are security core ontologies? A systematic mapping study.” In *15th International Conference on Research Challenges in Information Science*, edited by Samira Cherfi, Anna Perini, and Selmin Nurcan, 107–123. Limassol, Cyprus: Springer.

Oliveira, Ítalo, Tiago Prince Sales, Riccardo Baratella, Mattia Fumagalli, and Giancarlo Guizzardi. 2022. “An ontology of security from a risk treatment perspective.” In *International Conference on Conceptual Modeling*, edited by Jolita Ralyté, Sharma Chakravarthy, Mukesh Mohania, Manfred A Jeusfeld, and Kamalakar Karlapalem, 13607:365–379. Hyderabad, India: Springer.

Oltramari, Alessandro, Lorrie Faith Cranor, Robert J Walls, and Patrick D McDaniel. 2014. “Building an Ontology of Cyber Security.” In *Ninth Conference on Semantic Technology for Intelligence*, Defense, and Security, edited by Kathryn Blackmond Laskey, Ian Emmons, and Paulo C G Costa, 1304:54–61. Fairfax, Virginia, USA: ceur-ws.org.

Oltramari, Alessandro, Diane S Henshel, Mariana Cains, and Blaine Hoffman. 2015. “Towards a Human Factors Ontology for Cyber Security.” In *Tenth International Conference on Semantic Technology for Intelligence, Defense, and Security*, edited by Kathryn Blackmond Laskey, Ian Emmons, Paulo C G Costa, and Alessandro Oltramari, 1523:26–33. Fairfax, Virginia, USA: ceur-ws.org.

Otte, J Neil, John Beverley, and Alan Ruttenberg. 2022. “BFO: Basic Formal Ontology.” *Applied Ontology* 17 (1): 17–43.

Park, Hyunsoung, and Sangyun Shin. 2023. “A Proposal for Basic Formal Ontology for Knowledge Management in Building Information Modeling Domain.” *Applied Sciences* 13 (8): 4859.

Parr, Thomas, Giovanni Pezzulo, and Karl J Friston. 2022. *Active Inference: The Free Energy Principle in Mind, Brain, and Behavior*. Cambridge, Massachusetts, USA: The MIT Press.

Partridge, Chris, Andrew Mitchell, Al Cook, Jan Sullivan, and Matthew West. 2020. *A Survey of Top-Level Ontologies to inform the ontological choices for a Foundation Data Model*. Technical report. Cambridge, UK: Centre for Digital Built Britain, University of Cambridge.

Piccinini, Gualtiero. 2007. “Computing mechanisms.” *Philosophy of Science* 74 (4): 501–526.

Piccinini, Gualtiero. 2015. *Physical Computation: A Mechanistic Account*. Oxford, UK: Oxford University Press.

Piccinini, Gualtiero. (2018) Computational mechanisms. In Stuart Glennan & Phyllis M. Illari (Eds.), *The Routledge Handbook of Mechanisms and Mechanical Philosophy* (pp. 435–446). New York, New York, USA: Routledge.

Pritchard, Duncan. 2009. *Knowledge*. Basingstoke, England, UK: Palgrave Macmillan.

Rao, Rajesh P N, and Dana H Ballard. 1999. “Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects.” *Nature Neuroscience* 2 (1): 79–87.

Ray, Patrick L, Alexander P Cox, Mark Jensen, Travis Allen, William Duncan, and Alexander D Diehl. 2016. “Representing vision and blindness.” *Journal of Biomedical Semantics* 7 (Article 15): 1–12.

Reiersen, Jon. 2017. “Trust as belief or behavior?” *Review of Social Economy* 75 (4): 434–453.

Rodrigues, Fabrício Henrique, and Mara Abel. 2019. “What to consider about events: A survey on the ontology of occurrents.” *Applied Ontology* 14 (4): 343–378.

Rosa, Eugene A. 1998. “Metatheoretical foundations for post-normal risk.” *Journal of Risk Research* 1 (1): 15–44.

Sales, Tiago Prince, Fernanda Baião, Giancarlo Guizzardi, João Paulo A Almeida, Nicola Guarino, and John Mylopoulos. 2018. "The common ontology of value and risk." In *International Conference on Conceptual Modeling*, edited by Juan C Trujillo, Karen C Davis, Xiaoyong Du, Zhanhuai Li, Tok Wang Ling, Guoliang Li, and Mong Li Lee, 11157:121–135. Xian, China: Springer.

Scheuermann, Richard H, Werner Ceusters, and Barry Smith. 2009. "Toward an ontological treatment of disease and diagnosis." *Summit on Translational Bioinformatics* 2009:116–120.

Smart, Paul R. 2021. "Shedding Light on the Extended Mind: HoloLens, Holograms, and Internet-Extended Knowledge." *Frontiers in Psychology* 12 (Article 675184): 1–16.

Smart, Paul R, Wendy Hall, and Michael Boniface. 2022a. "Relativistic Conceptions of Trustworthiness: Implications for the Trustworthy Status of National Identification Systems." *Data & Policy* 4 (Article e21): 1–16.

Smart, Paul R, Wendy Hall, and Michael Boniface. 2022b. *Trust and Trustworthiness in the Internet of Things*. Technical report. Southampton, UK: Electronics and Computer Science, University of Southampton.

Smart, Paul R, Kieron O'Hara, Adrian Cox, and Wendy Hall. 2020. Cyber-Physical Systems and Social Machines. Available at SSRN: https://ssrn.com/abstract=3705252.

Smart, Paul R, Kieron O'Hara, and Wendy Hall. 2021. "Applying Mechanical Philosophy to Web Science: The Case of Social Machines." *European Journal for Philosophy of Science* 11 (3): 1– 29.

Smith, Barry, Farhad Ameri, Hyunmin Cheong, Dimitris Kiritsis, Dusan Sormaz, Chris Will, and J Neil Otte. 2019. "A First-Order Logic Formalization of the Industrial Ontologies Foundry Signature Using Basic Formal Ontology." In *10th International Workshop on Formal Ontologies meet Industry (FOMI)*, edited by Adrien Barton, Selja Seppälä, and Daniele Porello, 1–12. Graz Austria: CEUR-WS.org.

Smith, Barry, Michael Ashburner, Cornelius Rosse, Jonathan Bard, William Bug, Werner Ceusters, Louis J Goldberg, Karen Eilbeck, Amelia Ireland, and Christopher J Mungall. 2007. "The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration." *Nature Biotechnology* 25 (11): 1251–1255.

Smith, Barry, and Werner Ceusters. 2010. "Ontological realism: A methodology for coordinated evolution of scientific ontologies." *Applied Ontology* 5 (3–4): 139–188.

Smith, Barry, and Werner Ceusters. 2015. "Aboutness: Towards foundations for the information artifact ontology." In *Sixth International Conference on Biomedical Ontology*, edited by Francisco M Couto and Janna Hastings, 1515:1–5. Lisbon Portugal: CEUR Workshop Proceedings.

Smith, Barry, and Pierre Grenon. 2004. "The Cornucopia of Formal-Ontological Relations." *Dialectica* 58 (3): 279–296.

Smith, Barry, Tatiana Malyuta, Ron Rudnicki, William Mandrick, David Salmen, Peter Morosoff, Danielle K Duff, James Schoening, and Kesny Parent. 2013. "IAO-Intel: an ontology of information artifacts in the intelligence domain." In *Eighth International Conference on*

Semantic Technologies for Intelligence, Defense, and Security, edited by Kathryn Blackmond Laskey, Ian Emmons, and Paulo C G Costa, 1097:33–40. Fairfax, Virginia, USA: ceur-ws.org.

Spring, Jonathan M, and Eric Hatleback. 2017. "Thinking about intrusion kill chains as mechanisms." *Journal of Cybersecurity* 3 (3): 185–197.

Spring, Jonathan M, and Phyllis Illari. 2019. "Building general knowledge of mechanisms in information security." *Philosophy & Technology* 32 (4): 627–659.

Tchouanguem, Justine Flore, Mohamed Hedi Karray, Bernard Kamsu Foguem, Camille Magniont, F Henry Abanda, and Barry Smith. 2021. "BFO-based ontology enhancement to promote interoperability in BIM." *Applied Ontology* 16 (4): 453–479.

Toyoshima, Fumiaki. 2020. "Natural necessity: An introductory guide for ontologists." *Applied Ontology* 15 (1): 61–89.

Walls, Ramona L, Laurel Cooper, Justin Elser, Maria Alejandra Gandolfo, Christopher J Mungall, Barry Smith, Dennis W Stevenson, and Pankaj Jaiswal. 2019. "The plant ontology facilitates comparisons of plant development stages across species." *Frontiers in Plant Science* 10 (Article 631): 1–17.

Ward, Tony, and Claire Stewart. 2003. "Criminogenic needs and human needs: A theoretical model." *Psychology, Crime & Law* 9 (2): 125–143.

West, Matthew. 2011. *Developing High Quality Data Models*. Burlington, Massachusetts, USA: Morgan Kaufmann Publishers.

Wilson, Robert A., and Craver, Carl F. 2007. "Realization: Metaphysical and Scientific Perspectives." In Paul Thagard (Ed.), *Philosophy of Psychology and Cognitive Science* (pp. 81-104). Oxford, England, UK: North-Holland.

APPENDIX A: ACRONYMS AND ABBREVIATIONS

AI	Artificial Intelligence
BFO	Basic Formal Ontology
BIM	Building Information Modeling
C3O	Common Core Cyber Ontology
CCO	Common Core Ontologies
CIDO	Coronavirus Infectious Disease Ontology
CLIF	Common Logic Interchange Format
COVER	Common Ontology of Value and Risk
CPS	Cyber-Physical System
DBB	Digital Built Britain
DoS	Denial of Service
DOLCE	Descriptive Ontology for Linguistic and Cognitive Engineering
GFO	General Formal Ontology
HVAC	Heating, Ventilation, and Air Conditioning
HUFO	Human Factors Ontology
IAO	Information Artifact Ontology
IMF	Information Management Framework
IOF	Industrial Ontologies Foundry
IoT	Internet of Things
IoTSEC	Internet of Things Security Ontology
ISO	International Organization for Standardization
MAMO	MITRE ATT&CK Matrix Ontology
NDT	National Digital Twin
OBO	Open Biological and Biomedical Ontology
OWL	Web Ontology Language
RFID	Radio Frequency Identification
SOfIoTS	Secure Ontologies for the Internet of Things
SOSA	Sensor, Observation, Sample, and Actuator
SNOMED CT	Systematized Nomenclature of Medicine Clinical Terms
SSN	Semantic Sensor Network
ROSE	Reference Ontology for Security Engineering
TCP	Transmission Control Protocol
UCO	Unified Cyber Ontology
UFO	Unified Foundational Ontology
UML	Unified Modeling Language
VIDO	Virus Infectious Disease Ontology
YAMATO	Yet-another more advanced top-level ontology