Advances on Knowledge Representation Journal

Between Heaven and Earth: The Common Core of Basic Formal Ontology

John Beverley[1]

Abstract. *This paper critically examines the relationship between Basic Formal Ontology (BFO) and the Common Core Ontologies (CCO), offering a comprehensive discussion of their theoretical foundations, design patterns, and implementation practices. The paper highlights BFO's commitment to realism, perspectivalism, fallibilism, and adequatism, and illustrates how these principles guide the representation of domain-specific entities within CCO. The modular structure of CCO is analyzed, emphasizing its eleven component ontologies and the hub-and-spoke strategy that promotes semantic integration across diverse domains. The paper also engages with practical challenges in distinguishing between TLOs, MLOs, and domain ontologies, proposing heuristic and formal criteria for delineating their scope. Using a dataset from the US Federal Aviation Administration, the paper demonstrates CCO's modeling capabilities, particularly in integrating design specifications and real-world data. The study concludes by underscoring the importance of aligning data quality and semantic interoperability in ontology engineering, and it calls for sustained methodological rigor and collaboration to advance the BFO-CCO ecosystem.*

1. Introduction

The increasing complexity of data ecosystems, coupled with advances in artificial intelligence, has driven renewed interest in structured semantic representations capable of unifying heterogeneous data sources [1, 2]. *Ontologies* – logically well-defined controlled vocabularies representing entities and relationships among them [3] - have long been employed to serve this purpose across communities and systems. Ontologies have been leveraged to support data standardization, integration, machine learning, natural language processing, and automated reasoning [4] in fields such as biology and medicine [5], digital twins and IoT [6, 7], and proprietary AI systems [8]. If pursued without oversight, however, ontology engineering can easily recreate the semantic interoperability problems they are typically designed to address [9]. This occurs, for instance,

[1] Assistant Professor of Philosophy. Co-Director of the National Center for Ontological Research. Email: johnbeve@buffalo.edu. Affiliation: Institute for Artificial Intelligence and Data Science, University at Buffalo, NY, USA. ORCID: https://orcid.org/0000-0002-1118-1738.

when ontologies representing entities specific to one domain are created without concern for integration with nearby ontologies, resulting in *ontology silos*.

Silos often arise from incompatible data formats, divergent vocabularies, legacy systems, or even deliberate compartmentalization; silos hinder efficiency, decision-making, and innovation. Decades ago a desire to avoid the creation of ontology silos led to the creation of the first of several ontology "foundries" [10] aimed at standardizing ontology engineering practices. Among the principles underwriting most such efforts is that ontologies should extend[2] from a common top-level ontology (TLO). The most widely adopted TLO for such purposes is Basic Formal Ontology (BFO) [11], which defines the most general categories of being, such as **process**,[3] **material entity**, **object**, as well as relations such as **continuant part of**, **location of**, and so on.[4] BFO is an ISO/IEC 21838-2 TLO standard [12] used in over 700 open-source ontology initiatives, providing a semantically rich environment for data across biomedicine, manufacturing, defense and intelligence, and education, among others. BFO is small by design, allowing considerable flexibility for extension ontologies which use it as a starting point for ontology development. This flexibility is, however, a double-edged sword, since the engineering burden is then placed on the shoulders of those extending BFO.

Recognition of the need for guardrails that extend beyond the highly general classes and relations of BFO, led to the development of mid-level ontology (MLO) extensions, such as the Common Core Ontologies (CCO)[5] [3], which represents entities at a level more specific than BFO but more general than domain-level ontologies which more closely reflect data. CCO, for example, contains classes such as **agent**, **artifact function**, and **information content entity**, as well as relations such as **is about** and **is measurement of**. CCO is widely adopted across US defense and intelligence agencies owing both to the guardrails it provides for domain-level modeling as well as its unambiguous semantic connection to BFO [13, 14, 15, 16, 17]. As evidence, in 2024 both BFO and CCO were endorsed as "baseline standards" for ontology development across the US Department of Defense and Intelligence Community [18].

The success of the relationship between BFO and CCO can be understood on analogy with the successes of high-level programming languages and libraries which extend from them. BFO is analogous to languages such as Python, while CCO is analogous to libraries which extend such programming languages, such as NumPy or Pandas. Just as it is more sustainable to use an extension of Python when needed rather than build a library or language from scratch, it is more sustainable to leverage an extension of BFO when needed, rather than build an ontology from scratch. The former helps avoid the creation of ontology silos while the latter promotes their creation.

Adopting BFO and CCO does more than avoid ontology silos; it supports automated improvements to *data quality*—the fitness of data for its intended use. The low quality of data remains a persistent challenge that costs organizations billions annually in lost opportunities,

[2]Ontology *O* extends ontology *O** when *O** is a refinement of the intended interpretation of *O* achieved by adding new class vocabulary to *O*.
[3]We adopt the convention of displaying ontology terms and relations in bold. Pluralized term and relation names should be read as referencing instances. For example, "**processes**" is understood to mean "instances of **process**".
[4]BFO is under CC BY 4.0: https://github.com/BFO-ontology/BFO-2020
[5]CCO under the BSD-3: https://github.com/CommonCoreOntology/CommonCoreOntologies

remediation efforts, and incorrect decisions based on faulty guidance [19]. A key insight from ontology engineers in the BFO ecosystem is that *promoting interoperability and improving data quality are best pursued together*. Unfortunately, interoperability is sometimes sought without reflection on how data quality may be impacted and improving data quality is sometimes sought without reflection on how interoperability will be achieved [9, 21]. Too much emphasis on the former finds one too far from actual data to improve its quality without significant manual effort; too much emphasis on the latter leads back to ontology silos, unable to leverage common semantics in the name of interoperability. For example, if one tries to resolve interoperability challenges by simply stitching together ontology silos or heterogenous databases using case-by-case scripts, then as data evolves, is versioned, and datasets added, more case-by-case stitching will be needed, as well as re-checking and updating previous stitches.

Addressing interoperability and data quality challenges are best done in concert, and this observation motivates methodological principles underwriting BFO and CCO, as well as principles governing their implementations. In what follows, we defend the commitments and implementation practices of each, clarifying their respective roles in supporting interoperability and improved data quality. Our discussion will thus engage with underlying theories and architectural features of BFO and CCO, leading naturally to a defense of how best to draw the line between TLOs, MLOs, and domain-level ontologies that extend them. Additionally, we will examine how BFO and CCO facilitate ontology modeling, illustrating with a challenging real-world example using a dataset from the US Federal Aviation Administration. We close by discussing limitations and work to be done to sustain and mature the expanding BFO ecosystem.

2. Basic Formal Ontology

2.1 Principles of Basic Formal Ontology

BFO is distinguished from other TLOs by the adoption of four principles [9, 11]: *realism*, *perspectivalism*, *fallibilism*, and *adequatism*. **Realism** maintains that ontologies describe what exists in the world, independently of how humans think or talk about it. It is typically contrasted with *conceptualism*, which maintains that ontologies describe language or concepts.[6] Granted, many of the entities represented by ontologies will be human-made or dependent on human activities for their existence: this is the case, for example, for games of chess, moral systems, novels, and so on. Yet, these entities and their dependencies exist independently from the way in which they are conceived or spoken about in everyday language. As such, they are within the scope of BFO's realism. In other words, BFO's realism is broad enough to encompass conceptualism thus understood.

[6] Conceptualism is often motivated by citing Gruber's definition of an ontology as "a specification of a conceptualization." [21, 22] Less often noted, however, is that Gruber's definition was accompanied by a citation to Genesereth and Nilsson's 1987 *Logical Foundations of Artificial Intelligence*, which contains an entire chapter on "conceptualization", described therein as abstract representations that may be about, among other things, "objects presumed or hypothesized to exist in the world and their interrelationships." [23] Gruber's definition is neutral between realism and conceptualism. See [24] for an interesting discussion.

While reality exists independently from our perspectives of it, it is also so complex that no single perspective can justifiably claim to exhaust it. Different disciplines and domains often require different and yet equally good representations of the same entities, as evidenced by for example the use of "species" language across biological and ecological sciences [25]. To accommodate, BFO adopts **perspectivalism**, ontologies should aim at equally representing all warranted outlooks on reality, without necessarily trying to create one grand unified theory that fits them all together []. This commitment has ramifications *within* domains just as it does *across* them. For example, the International Committee on Taxonomy of Viruses (ICTV) is a very deep, very extensive, taxonomy widely used by virologists [26, 27]. There is, however, a more parsimonious way to classify viruses based on the seven most common replication pathways grounded in viral genetic structure, which is also leveraged by virologists [28]. Representing each perspective on the domain of viruses within the scope of BFO-conformant ontologies is both justified and desirable.

Fallibilism accepts that our best understanding of reality is subject to revision; hence, ontologies must be modifiable in light of new evidence. As human knowledge grows and matures, ontologies representing that knowledge must grow and mature as well. Terms and relations within BFO ontologies are thus subject to deletion, correction, update and extension and so should be developed with corresponding tools and practices that support these processes. An illustrative case involved the International Astronomical Union (IAU) working definition of "planet" as a celestial body whose mass exceeds a certain threshold. Under this working definition, Pluto was considered a planet. IAU members resolved in 2006 that for a celestial body to count as a planet, it must orbit the Sun, have sufficient mass to maintain hydrostatic equilibrium, and be the dominate gravitational force in its orbit [29, 30]. Since Pluto's orbit is determined by Neptune's gravitational field, it no longer counted as a planet [30]. Because BFO accepts changes in the scientific understanding of the world as reasons to revise ontological commitments, it is not enough here to simply say Pluto was a planet at some past time and not presently; rather, Pluto never was a planet, despite being miscategorized as one. When the IAU later clarified the definition of planet, they are stating what has always been true about our world.

Adequatism maintains that ontologists should refrain from trying to reduce terms and relations to others, assuming the entities reflected are taken seriously by researchers within some domain. This means that not only we have to recognize the existence of, say, cats, organs and subatomic particles; we also should refrain from reducing cats to collections of organs and organs to collections of particles. Adequatism is closely tied to the notion of *granular partitions*—the idea that we classify and divide the world into categories according to different levels of detail or resolution [31]. Different granularities allow us to focus on different aspects of reality, emphasizing certain features while ignoring others. For example, in studying an organism, an anatomical partition highlights organs and tissues, while a chemical partition highlights molecules and bonds. Each partition brings out different aspects of the same underlying entity, and both are essential for a complete understanding of reality.

2.2 Basic Formal Ontology

Terms in BFO and in BFO-conformant ontologies represent classes of instances that share important features. The highest division in BFO's class taxonomy is between **occurrent** and **continuant**. **Occurrents** are extended in time in such a way as to have temporal parts whereas **continuants** lack temporal parts and endure through time. **Continuants** and **occurrents** are tied together by the fact that the former **participate in** the latter, as when a child **participates in** an act of crying or a mother **participates in** an act of consoling.

There are three subclasses of **continuant**, illustrated in **Figure 1**. An **independent continuant** is a **continuant** that does not depend on anything for its existence [9].[7] A landmass is an **independent continuant**, its mass and shape, on the other hand, depend for their existence on the landmass and are categorized as **specifically dependent continuants**, which depend for their existence on some specific **independent continuant**.

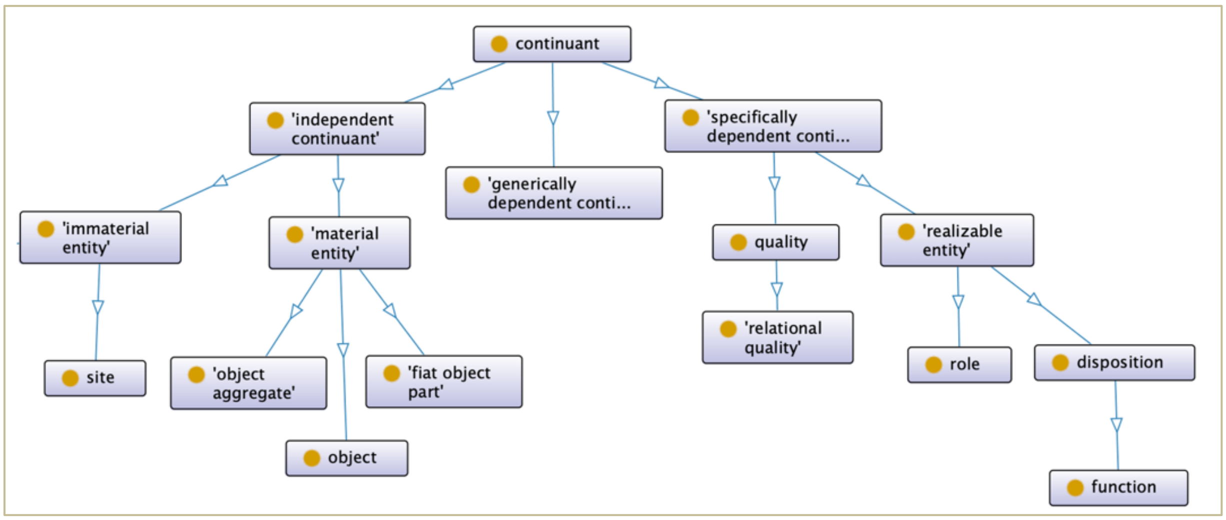

Figure 1: Basic Formal Ontology Continuant Taxonomy

Independent continuant has two BFO sub-classes: **material entity** and **immaterial entity**, the former having and the latter lacking material parts. Subclasses of **material entity** include: **objects**, such as Taylor Swift, **object aggregates**, such as the Destiny's Child, and **fiat object parts,** such as the left hemisphere of Beyonce's brain. **Object aggregates** consist of disjoint unions of **objects**, whereas **objects** consist of **material entities** that are maximal with respect to some causal unity criterion, such as the unity exhibited by surface boundaries, forces between fundamental particles, or fastening via engineering processes. Fiat object parts are in turn mereological parts of **objects** that warrant their own class, such as your head or the Southern Hemisphere of the Earth. In each case, the **material entity** class is closed under BFO's

[7] *Depends on* holds between *x* and *y* when the former is such that it cannot exist unless the latter exists [32].

mereological **continuant part of** relation, such that any entity that has a **material entity** part is a **material entity**. Subclasses of **immaterial entity** include **spatial regions** and **sites** (for example a hole in the ground on a golf course), as well as various continuant boundary entities (for example the boundary separating a golf course from a highway).

Certain instances of **specifically dependent continuant** are fully manifested whenever they manifest at all, such as color, shape, or mass; these are instances of the class **quality**. **Realizable entities**, in contrast are those **specifically dependent continuants** which are marked by the fact that they may exist without manifesting. For example, if a flotation device can float in water, it possesses this ability even when it is not deployed. **Realizable entities** must be realiz*able*, but they are not in every case realiz*ed*. Two major subclasses of **realizable entity** recognized by BFO are **dispositions** and **roles**. **Dispositions** are **realizable entities** that are *internally grounded*, which means for a **disposition** to begin or cease to exist, its bearer must undergo a physical change. For example, a portion of salt may lose its solubility but only if it undergoes some physical change to its physical structure [33]. **Role** is a disjoint sibling class of **disposition** whose instances are optional in the sense that bearers may gain or lose them without thereby exhibiting material change. They are, moreover, *externally grounded*. A student who graduates from a university no longer bears the **role** of student at that institution, but that need not entail any change to the physical structure of the student. This feature allows **roles** to be borne by entities that do not have material parts, such as the boundaries of a country, the location where a river used to be, or the internal cavity of a bear's mouth. **Dispositions** are not afforded such a status, given their dependence on the material structure of bearers. The class **disposition** has a single subclass, namely **function,** which is a **disposition** that reflects the reason for the existence of its bearer, such as the heart's **function** to pump blood or the **function** of a knife to cut. In each case, the reason the bearer exists is because it has the **function** it bears.

Generically dependent continuant is a sibling class of **independent continuant** and **specifically dependent continuant**. A **generically dependent continuant** is – roughly - a copyable pattern. A pattern exists only if it is **concretized in** some bearer; but it is not dependent on any specific bearer, because it may be copied (for example through being transmitted) from one bearer to another. As we will see in our example below, **generically dependent continuant** plays an important role when modeling with CCO.

AKR Advances on Knowledge Representation Journal

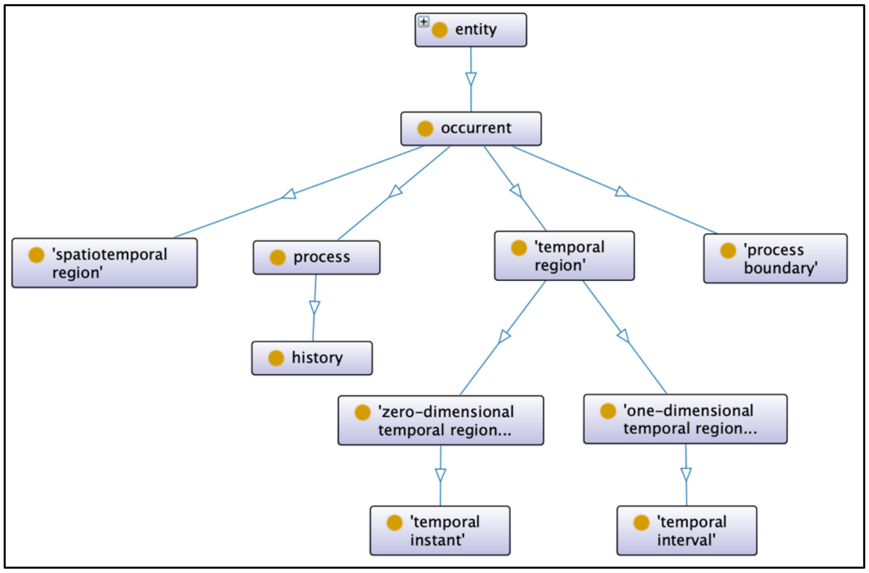

Figure 2: Basic Formal Ontology Occurrent Taxonomy

Figure 2 displays the BFO **occurrent** hierarchy, subclasses of which cover time, processes, boundaries of processes, and spacetime. Instances of **process** have temporal parts and must have at some point a **material entity** which **participates in** them. For example, baking a cake is a **process** which might involve **material entity** ingredients which **participate in** the **process** insofar as they are introduced and consumed by it, once a cake is produced. **History** is a subclass of **process** reflecting the totality of all **processes** in which a given **material entity participates**. It is assumed, moreover, that each **material entity** has exactly one **history**. Poor John may lose his hand one evening, at which point the **history** of John, and the **history** of John's hand diverges. Such **processes** would have boundaries, a beginning and an end, instances of which would fall under the BFO class **process boundary**. Both **processes** and **process boundaries** will unravel over some **temporal region**, the former over a **temporal interval** while the latter at a **temporal instant**.[8] Where BFO's **participates in** relationship bridges **processes** and **continuants**, the **spatiotemporal region** class bridges the underlying **spatial regions** and **temporal regions** in and over which **continuants** exist. **Continuants participate in processes** which exist in **spatiotemporal regions**, which **project on** the **spatial regions** in which the relevant **continuant** is **located** and the **temporal region** over which the relevant **process** occurs. **Table 1** provides further descriptions of the BFO hierarchy.

[8]**Temporal interval** and **instant** are subclasses of – rather than equivalent to – **one** and **zero-dimensional temporal region**, respectively. An aggregate of disconnected **temporal intervals** is not a **temporal interval** but is a **one-dimensional temporal region**. Similarly, an aggregate of disconnected **temporal instants** is not a **temporal instant** but is a **zero-dimensional temporal region**.

Adv. Know. Repr. 2025; volume 5; number: 3. DOI 10.5281/zenodo.17209739.
Special Issue on Top-Level and Middle-Level Ontologies.

AKR Advances on Knowledge Representation Journal

BFO Class	Elucidation[9]/Definition
continuant	An entity that persists, endures, or continues to exist through time while maintaining its identity.
independent continuant	A continuant which is such that there is no *x* such that it specifically depends on *x* and no *y* such that it generically depends on *y*.
specifically dependent continuant	A continuant which is such that (i) there is some independent continuant *x* that is not a spatial region, and which (ii) specifically depends on *x*.
generically dependent continuant	An entity that exists in virtue of the fact that there is at least one of what may be multiple copies.
material entity	An independent continuant that at all times at which it exists has some portion of matter as continuant part.
object	A material entity which manifests causal unity and is of a type instances of which are maximal relative to the sort of causal unity manifested.
object aggregate	A material entity consisting exactly of a plurality (≥1) of objects as member parts which together form a unit.
quality	A specifically dependent continuant that, in contrast to roles and dispositions, does not require any further process in order to be realized.
realizable entity	A specifically dependent continuant that inheres in some independent continuant which is not a spatial region and is of a type some instances of which are realized in processes of a correlated type.
role	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 is not such that, if it ceases to exist, then the physical make-up of the bearer is thereby changed.
disposition	A realizable entity such that if it ceases to exist, then its bearer is physically changed, and its realization occurs when and because this bearer is in some special physical circumstances, and this realization occurs in virtue of the bearer's physical make-up.
function	A disposition that exists in virtue of the bearer's physical make-up and this physical make-up is something the bearer possesses because it came into being, either through evolution (in the case of natural biological entities) or through intentional design (in the case of artefacts), in order to realize processes of a certain sort.
occurrent	An entity that unfolds itself in time or is the start or end of such an entity or is a temporal or spatiotemporal region.
process	An occurrent that has some temporal proper part and for some time has a material entity as participant.
x generically depends on y	x is a generically dependent continuant & y is an independent continuant that is not a spatial region & at some time t there inheres in y a specifically dependent continuant which concretizes x at t
x continuant part of y	x and y are continuants & there is some time t such that x and y exist at t & x continuant part of y at t
x occurrent part of y	A relation between occurrents x and y when x is part of y

[9]Elucidations are descriptions provided to help fix the referent of primitive terms; definitions express individually necessary and jointly sufficient conditions for an entity to be an instance of the class defined.

x participates in y	participates in holds between some x that is either a specifically dependent continuant or generically dependent continuant or independent continuant that is not a spatial region & some process y such that x participates in y some way
x concretizes y	x is a process or a specifically dependent continuant & y is a generically dependent continuant & there is some time t such that y is the pattern or content which x shares at t with actual or potential copies
x inheres in y	x is a specifically dependent continuant & y is an independent continuant that is not a spatial region & x specifically depends on y

Table 1: Definitions/Elucidations of Selected BFO Content

3. Common Core Ontologies

3.1 Principles of the Common Core Ontologies

CCO adopts the principles of BFO described above but also highlights implementation commitments common in the BFO ecosystem, in particular adherence to *ontology modularization* and the *hub-and-spoke* strategy. Ontology modules are standardly characterized as self-contained components of ontologies, often designed to be integrated with other self-contained components of ontologies [34]. The CCO suite exhibits this sort of modularization with self-contained components scoped to temporal, geospatial, and informational domains, among others. Modularization provides flexibility when deploying CCO into applications, as not every use case needs the full expressivity of the suite.

Ontology modules, moreover, provide the foundation for the hub-and-spoke strategy, where *ontology hubs* are ontology modules designed to serve as foundations from which more specific ontologies – *ontology spokes* – extend [35]. The CCO modules listed in **Table 2** are ontology hubs, and serve as foundations for CCO ontology spokes, such as the Space Object Ontology [16]. Complementing the hub-and-spoke strategy is the preservation semantic continuity through spokes and hubs ultimately back to BFO by insisting that terms and relations be defined following the definition schema: "A is a B that C's" where "A" is the term being defined, "B" is its immediate parent in the subclass hierarchy, and "C's" are whatever attributes distinguish A from other subclasses of B [9, 11]. For example, the CCO class **agent** which stands atop the Agent Ontology is defined with respect to its parent in BFO: **material entity**, differentiated by other **material entities** insofar as instances may engage in intentional actions.

Module	Scope
Geospatial Ontology	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.
Information Entity Ontology	Designed to represent generic types of information as well as the relationships between information and other entities
Event Ontology	Designed to represent processual entities, especially those performed by agents, that occur within multiple domains.
Time Ontology	Designed to represent temporal regions and the relations that hold between them.
Agent Ontology	Designed to represent agents, especially persons and organizations, and their roles.

Advances on Knowledge Representation Journal

Quality Ontology	Designed to represent a range of attributes of entities especially qualities, realizable entities, and process profiles.
Units of Measure Ontology	Designed to represent standard measurement units that are used when measuring various attributes of entities.
Currency Unit Ontology	Designed to represent currencies that are issued and used by countries.
Facility Ontology	Designed to represent buildings and campuses that are designed to serve some specific purpose, and which are common to multiple domains.
Artifact Ontology	Designed to represent artifacts that are common to multiple domains along with their models, specifications, and functions.
Extended Relations Ontology[10]	Designed to represent many of the relations that hold between entities at the level of the mid-level Common Core Ontologies.

Table 2: Eleven Modules of the Common Core Ontology Suite

3.2 Common Core Ontologies Key Patterns

Compared to BFO, CCO is quite large. Rather than introduce classes and relations through each of its hubs, we will instead explore the content of through key design patterns reflected in its modules.

The Information Entity Ontology focuses on representing the types and provenance of information. A major design pattern distinguishes between information - **information content entities** - its physical carriers - **information bearing entities** – and the **specifically dependent continuants** that **concretize** that information. For example, the **qualities** of your computer screen and mine may **concretize** the same **information content entity** associated with a given PDF, with each computer hard drive being an **information bearing entity**. Another key design pattern is the primitive **is about** relationship [36], which connects an **information content entity** to an **entity** that it is about. This relationship is specialized further with respect how the relevant information content entity stands in this aboutness relation. The information content of a newspaper article **describes** some current event, much like an accident report **describes** some accident. A blueprint **prescribes** a model for some product, much like a professional code of conduct **prescribes** a set of rules for anyone acting in a professional role. The content of a photograph **represents** the photographed entity, much like the content of a transcript **represents** the verbal interaction transcribed. The sense of "isomorphism" in the definition of **represents** is understood relative to the type of entities involved. For example, the arrangement of Napoleon's body parts in a painting by Jacques Louis David was meant to reflect the actual arrangement of Napoleon's body.

The Agent Ontology models entities that can act intentionally—**persons** and **organizations**. It distinguishes between **agent qualities** such as height or weight and the **roles** they may assume in certain contexts, such as citizen or manager. **Agents** may thus **participate in processes**, bear **roles** and **qualities**. A key design pattern is the role-bearer model, where, for example, specializations of agents are defined in terms of **roles**. For example, a **citizen** is defined as a **person** bearing a **citizen role**. Closely related is the Event Ontology provides terms and relations for modeling **acts** and **processes**, along with the **agents** and **objects** that **participate in**

[10]The Extended Relations Ontology is based on the Relations Ontology of the Open Biological and Biomedical Ontologies Foundry [10].

Adv. Know. Repr. 2025; volume 5; number: 3. DOI 10.5281/zenodo.17209739.
Special Issue on Top-Level and Middle-Level Ontologies.

them. Its major design pattern is the event-participation model, which connects **objects** to **processes** and **processes** to **temporal** and **spatial regions**. Additional patterns include modeling **change** in attributes through change events, such as the gaining or losing of a role, and the use of **stasis** to capture continuous states, allowing for representations of both dynamic and steady-state conditions.

In the Artifact Ontology, the dominant pattern is the artifact-function design, where **artifacts** are defined in terms of assigned **functions**, such as a grenade bearing a **function** to detonate, or a knife bearing a **function** to cut. This pattern connects closely to design specifications and blueprints that are properly within the scope of the Information Entity Ontology. Simply put, artifact specifications **prescribe** the creation of **artifacts** bearing **functions** so prescribed. Specifications often also **prescribe artifacts** bear certain **qualities**, are found within the scope of the Quality Ontology which contains terms for a variety of **qualities**, **dispositions**, **roles**, and **process profiles**. **Qualities** and realizable entities behave as in BFO, and include **specifically dependent continuants** such as color, weight, or shape. The **process profile** design pattern supports representing dynamic attributes of **processes**, such as cyclical rates of change or temperature over time. These patterns together support granular modeling of both static and dynamic properties of entities.

The Geospatial Ontology applies a spatial containment pattern to locations like cities, states, and countries, related via **located in** and **location of**, as well as an implementation of the Regional Connection Calculus 8 (RCC8) relations, such as **externally connects with** and **overlaps with** [37]. The Geospatial Ontology thus supports formal spatial reasoning and integration across a variety of geospatial datasets. The Time Ontology supports temporal reasoning through relations like **interval overlaps**, **interval meets**, and **interval starts**, allowing for This design processes to be temporally aligned, compared, and integrated even when described at different temporal granularities.

In the Units of Measure Ontology, the key pattern is the separation of **measurement units** as individuals – for example, Fahrenheit unit is an instance - distinct from the measurements themselves, allowing for unit standardization and conversion. **Measurement units** are linked to **information bearing entities** that carry quantitative data while a measurement scale hierarchy is introduced for sorting scale types cleanly, as **ratio**, **interval**, **ordinal**, and **nominal measurement content entities**. The Currency Unit Ontology extends the Units of Measure Ontology pattern by specializing it to monetary values, treating currencies as **measurement units** of financial value.

The Extended Relation Ontology adopts a relation enrichment pattern, expanding the vocabulary from the Relations Ontology [38]. It defines additional object properties such as **has input** and **has output**, as well as the information-specific relations introduced above. It also incorporates metadata annotation properties for **definition**, **definition source**, and so on, that support documentation, provenance tracking, and querying within RDF-based systems. Each of these ontologies adheres to the BFO principles described earlier, grounding classes and relations within the BFO framework.

AKR Advances on Knowledge Representation Journal

3.3 BFO & CCO Modeling Scenario

Given the complexity of CCO, it is worth working through a specific modeling exercise to better illustrate its strength. For this exercise, we will use publicly accessible data from the US Federal Aviation Administration concerning the "essential characteristics of aircraft types…in order to perform airport planning and design function."[11] Using this data, we will model the following:

> *The Airbus A321-111 is designed to have a maximum know approach speed of 142 km/hr. However, after five approaches, a specific Airbus A321-111 has obtained an average knot approach speed of only 139.*

Perusal of the data will reveal that while the maximum knot speed is provided, there is no mention of five approaches and corresponding measurements. Still, the remainder of the scenario is plausible enough, and will highlight modeling nuance of both BFO and CCO. **Table 3** describes CCO definitions useful for modeling this scenario.

Label	Definition
information content entity	A generically dependent continuant that generically depends on some information bearing entity & stands in relation of aboutness to some entity
jet specification	An information content entity that prescribes the creation of a jet.
artifact	A material antity that was designed by some agent to realize a certain function.
artifact function	A function that inheres in some artifact in virtue of that artifact being designed to be used in processes that require that function to be realized.
motion artifact function	An artifact function that is realized in a process in which an entity changes its position with respect to time.
propulsion artifact function	A motion artifact function that is realized in a process in which the bearer of the function creates force leading to an entity's movement.
information bearing entity	Object upon which an information content entity generically depends
x represents y	x is an instance of information content entity, y is an instance of entity, & z is carrier of x & x is about y in virtue of there existing an isomorphism between characteristics of z & y
x describes y	x is an instance of information content entity & y is an instance of entity & x is about the characteristics by which y can be recognized or visualized
x prescribes y	x is an instance of information content entity & y is an instance of entity & x serves as a rule or guide for y if y an occurrent, or x serves as a model for y if y is a continuant
has manufacturer	A datatype property that relates an artifact to the literal reflecting its manufacturer.
has model FAA	A datatype property that relates an artifact to the literal reflecting its Federal Aviation Administration model number.
has maximum approach speed	A datatype property that relates an artifact to the maximum approach speed as defined by the Federal Aviation Administration.

Table 3: CCO Content and Additions Relevant to the Airbus Scenario

[11] https://www.faa.gov/airports/engineering/aircraft_char_database

Adv. Know. Repr. 2025; volume 5; number: 3. DOI 10.5281/zenodo.17209739.
Special Issue on Top-Level and Middle-Level Ontologies.

We first note that we are here dealing with a portion of an aircraft specification, which concerns how the world should be rather than how it is. Many have claimed modeling such "modal" phenomena in BFO is challenging owing to its commitment to realism. The sentiment seems to be that BFO-conformant ontologies must be designed to reflect reality as it is rather than as it could be, evidenced by modality in BFO relegated to **realizable entities** borne by existing entities, which may or may not be realized, where realizations are only represented if they occur. This purported limitation of BFO is, however, mistaken. Appealing to the Information Entity Ontology extension, we have already encountered how BFO-conformant ontologies may represent specifications for entities that do not presently exist. What holds for specifications holds equally for fictional entities, blueprints, simulations, and so on. The commitment adopted in BFO and CCO is simply that when representing such entities, refrain from introducing instances that do not exist. One can represent an Airbus specification that prescribes the creation of an instance of jet, but one should not create an instance of jet this specification is about until it rolls off the production line.[12]

An Airbus 321-111 specification is intended to prescribe possible arrangements of classes and relationships among them. This specification is a plan for an aircraft, but not about any specific instance that might emerge from production; rather, prescribe arrangements of portions of rubber and metal, properties of shape, size, and thermal conductivity, relations of parthood and dependence, and so on. The prescription exhibited by specifications aim at the class-level rather than instance-level. More specifically, we leverage the **prescribes** relation alongside a subclass of **information content entity** we introduce called **jet specification**:

- Airbus 321-111 specification instance of **jet specification** and *prescribes* only
    - **subclass of** some **fixed wing** and
    - **has continuant part** some **jet engine** and
    - **bearer of** some **propulsion artifact function** and
    - **has manufacturer** value Airbus and
    - **has model FAA** value Airbus A321-111
    - **has maximum approach speed** value 142
    - …

We leave as an exercise for the reader examining the column headers in the data set to see how our constraints compare. That aside, this characterization allows us to keep separate the Airbus specification from what it is about. The specification prescribes only something that is a member of a class that is a subclass of fixed wing, having expected continuant parts, bearing a disposition to transport, having a maximum approach speed, and so on. Once such a prescribed instance is produced, it will fall within this intersection; until then, the intersection is empty. Again, the prescription exhibited by specifications aim at the class-level rather than instance-level.

[12] It is worth noting that previous versions of CCO did as much in the Modal Relations Ontology [48], which introduced a **modal object property** under which duplicates of all CCO relations fell. Users were directed to model possibilities using this relation and possible instances.

The scenario assumes such an instance was produced, as its maximum approach speed is measured at least five times. **Figure 3** illustrates the specification and part of the measurements.

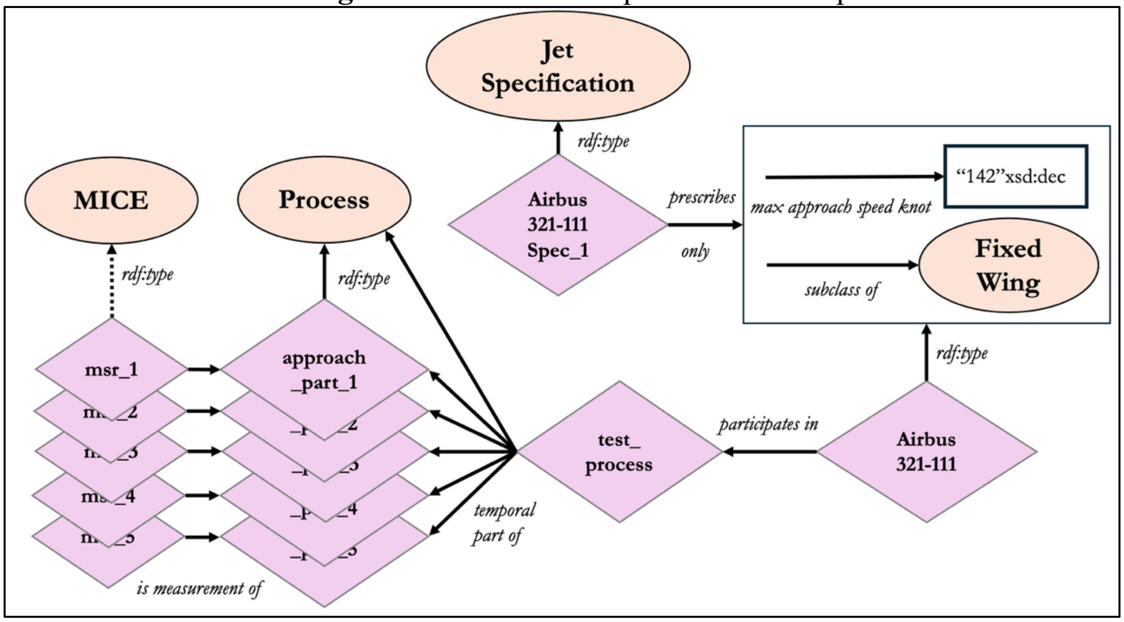

Figure 3: Airbus 321-111 Specification and Five Measurements; dotted arrows indicate shortcuts.

There is some Airbus 321-111 instance that **participates in** some **test process**, which has at least five temporal parts aligned in some precedence order. Each of these temporal parts, moreover, has the Airbus 321-111 instance as a **participant** and for each there is a **measurement information content entity** that **is about** the Airbus instance's approach speed over the corresponding **temporal interval**. Each **measurement information content entity** is then associated with some integer value as well as a knot unit instance. Lastly, there is an average that takes as input each of the five **measurement content entities**, which **is about** the entire test process instance. These five measurements, when averaged, result in "139", which is the **measurement value** of the average instance. **Figure 4** illustrates the remainder of the scenario as described.

AKR Advances on Knowledge Representation Journal

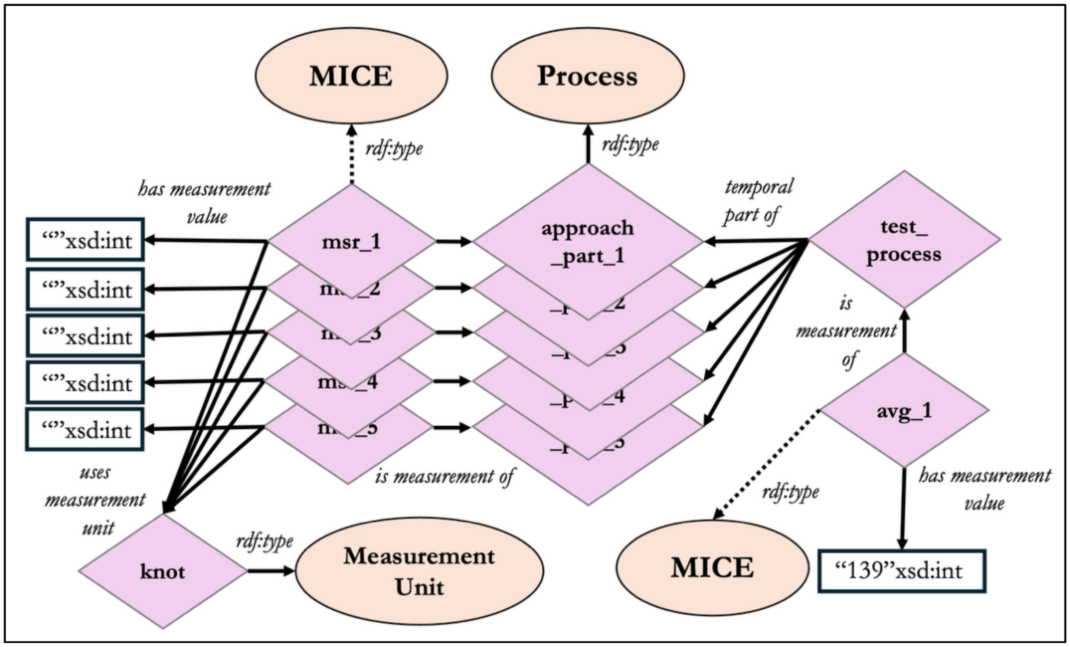

Figure 3: Airbus 321-111 Measurements and Average; dotted arrows indicate shortcuts.

3. Drawing Lines between TLOs, MLOs, and Extensions

The distinction between TLOs from MLOs is often characterized by levels of generality. TLOs are the most general whereas MLOs are considered more specific than TLOs but less specific than domain-level extension ontologies [39]. Neither characterization provides a firm cutoff, however, between TLOs and MLOs or between MLOs and their extensions. Little attention has been paid to any putative criteria for justifying a cutoff between TLOs and MLOs. Nevertheless, a useful heuristic when determining where to draw the line concerns ontology elements needed for a complete representation of a given domain. That is, when considering whether an ontology term or relation should be in a TLO or is better suited to an extension, observe that any complete representation of a given domain should leverage every term and relation in the TLO. So, if that term or relation is necessary for a complete representation of any domain, then it likely deserves a spot in the TLO. If not, then not. For example, a complete representation of the domain of chemistry will require leveraging molecular **material entities** (molecules), **sites** they inhabit, **qualities**, **roles** and **functions** they bear, relevant copyable patterns, transformations in which they **participate**, **boundaries** which they encounter, within space, over time, and so on. In contrast, complete representations of molecular entities need not involve **currency units**, **artifacts**, or **facilities**, among other entities one finds in CCO.

Much more discussion has been given to potential cutoff principles between MLOs and their extensions. Even so, it is in general unclear how best to identify a justifiable division between what

AKR Advances on Knowledge Representation Journal

should be included in an MLO rather than a domain ontology, or vice versa. A natural suggestion one finds discussed can be represented as [40]:

> **(*)** For a given ontology element *e*, natural number $n > 1$, and distinct domain-level ontologies $o_1 \ldots o_n$: If *e* is appropriately reused in $o_1 \ldots o_n$ then the primary residence of *e* should be a more general ontology imported by $o_1 \ldots o_n$.

(*) is, in certain circumstances, a useful principle. Consider that the term "infection" is plausibly used across all infectious disease ontologies. Housing **infection** in, say, an ontology whose scope is restricted to influenza would require other infectious disease ontologies to import **infection** from that influenza ontology. Better to place **infection** in a more general ontology alongside terms commonly used across multiple infectious disease domain ontologies. **(*)** justifies such a decision.

Unfortunately, because domain ontologies that extend from CCO may legitimately represent the same domain in different ways, **(*)** fails to provide a defensible cutoff between CCO and its domain extensions. Consider that a domain ontology intended to represent car accidents represents a domain that plausibly overlaps with the car insurance domain just as well as do domain ontologies built specifically to represent car insurance. Both ontologies may plausibly include a class **Hyundai Elantra** but this should not entail that **Hyundai Elantra** is a class that belongs in CCO. Similarly, a domain ontology representing strategies for recycling vehicles might also have need for **Hyundai Elantra** within its scope. But three domain ontologies using **Hyundai Elantra** should not force this class into CCO. One might still be tempted to claim that for some sufficiently large *n*, reuse across *n* domain ontologies warrants inclusion in CCO. However, because CCO adopts perspectivalism, it can be extended by overlapping but distinct domain-level ontologies in potentially infinite ways; thus, leveraging **(*)** – even for some large *n* – to provide a firm cutoff between CCO and domain ontologies runs the risk of collapsing them.

While rules of thumb have been suggested to draw a line between CCO and its extensions – such as limiting the number of its subclasses to no more than three [40] – associated rules are unhelpfully arbitrary. More promising is to rely instead on existing consensus regarding CCO content, as there is often more agreement as to what should be included than there is disagreement. For every contentious, potentially borderline class or relation in CCO implementations – such as **flywheel** or **is first cousin of** – there are many uncontentious classes – such as **agent**, **information content entity**, **measurement**, **is about**, and so on [3].

There are nevertheless defensible constraints should require of MLOs in general and CCO in particular. Any MLO should extend from some TLO and indeed an ISO/IEC 21838-1 TLO, such as BFO [12]. Following [39], we codify this principle as:

> **EXTEND** MLOs must extend from at least one ontology satisfying the requirements specified in ISO/IEC 21838-1.

This ensures that MLOs are in fact "middle" with respect to a vetted TLO. Additionally, MLOs should be composed of all and only ontology hubs, none of which are TLOs. Such a constraint is

Adv. Know. Repr. 2025; volume 5; number: 3. DOI 10.5281/zenodo.17209739.
Special Issue on Top-Level and Middle-Level Ontologies.

intended to exclude putative MLOs that are generated simply out of TLOs and domain-level ontologies that are better suited as ontology spokes, reflecting the intuition that MLOs are more general than domain-level ontologies but less general than TLOs:

> **HUB** MLOs should be composed of all and only ontology hubs none of which overlap in scope with any other.[13]

As a limit case, **HUB** can be satisfied by a single ontology hub. More generally, **HUB** may be satisfied by a collection of one or more ontology hubs. For example, as illustrated in **Table 2**, CCO is composed of eleven ontology hubs with distinct scopes.

MLOs should exhibit a tight connection with the TLOs which they extend. One way to achieve this is by requiring MLOs exhaust the scope of their TLO. Such a commitment conflicts, however, with certain characterizations of MLOs as ontologies “that represent relatively general categories common to many domains of interest.” [15] One way to interpret this characterization is to understand MLOs as picking out ontologies representing some broad user community or perhaps scientific field, such as biomedicine, manufacturing, education, and so forth. On such a picture, a given TLO might be extended by both a biomedical MLO, a distinct manufacturing MLO, a distinct education MLO, and so on. Call these *relative mid-level ontologies*.

Relative MLOs encourage *scope creep* [9], which arises when an ontology intended to represent some specific domain is constructed with insufficient foresight for possible extensions, so that it later needs to be expanded beyond that domain. Consider the Industrial Ontologies Foundry Core (IOFC), described by its developers as an MLO with respect to industrial manufacturing and services [41]. IOFC extends directly from BFO and so inherits only its minimal terms and relational expressions. Accordingly, IOFC developers found a need to mint new ontology vocabulary representing agents, artifacts, information, and so on, much of which was outside the scope of IOFC proper. This is scope creep. A natural antidote would be to store relevant terms and relational expressions representing artifacts, information, etc. needed by the IOFC relative MLO in a ‘more general’ MLO which IOFC imports. Scope creep is, however, pervasive among relative MLOs [42]. A natural progression of enough relative MLOs aiming to avoid scope creep would be to create a ‘most general MLO’, which is to say an MLO that exhausts the scope of the TLO from which it extends. More simply, coupling relative MLOs with pressure to avoid scope creep leads naturally to some MLO that exhausts the scope of its TLO. Scope creep is notoriously challenging to address once established. We should then encourage *starting* with such a ‘most general MLO’.

Adding to this, MLOs should inherit the scope of their TLO by introducing more specific ontology content. As a first pass:

> **(**)** MLOs must contain at least one subclass for each class reflecting the lowest level terms or relations of the TLO from which they extend.

[10] Compare [10] where it is argued that OBO Foundry ontologies should have orthogonal scope.

AKR Advances on Knowledge Representation Journal

For example, BFO classes such as **function** and **history** are extended in CCO to **artifact function** and **artifact history**, respectively. While **(**)** seems initially attractive, it is revealed on reflection to be too strong. Consider that subclasses of BFO's **spatial region** are still rarely, if ever, introduced correctly [43]. CCO currently includes subclasses for **one-dimensional spatial region** such as **Coordinate System Axis**, which is a "A One-Dimensional Spatial Region defined by a Coordinate System for the purpose of identifying the position of entities along one dimension of the Coordinate System's spatial framework." This definition strongly suggests the class should be a species of **generically dependent continuant**, as coordinate systems and their parts are copyable mathematical patterns. This is common among subclasses of BFO's **spatial region**. Extensions of child classes of 'spatial region' will, we believe, not be needed by most ontologies; it is plausible that some 21838-1 TLOs will include classes that should not be extended by MLOs. Hence, **(**)** is too strong.

21838-1 provides a way forward. Any top-level ontology satisfying this standard must provide explanations for how data across the breadth areas in **Figure 5** will be represented.

Space and Time	Qualities and other Attributes
Actuality and Possibility	Quantities and Mathematical Entities
Classes and Types	Processes and Events
Time and Change	Constitution
Parts, Wholes, Unity, Boundaries	Causality
Space and Place	Information and Reference
Scale and Granularity	Artifacts, Socially Constructed Entities
Mental entities, imagined entities, fiction, mythology, and religion	

Figure 5: ISO/IEC 21838-1 Coverage Areas

We may leverage these breadth areas to provide a constraint on MLOs more flexible than **(**)**:

INHERITANCE MLOs are composed of all and only content extended from each breadth area of the TLO referenced in **EXTEND**.

By requiring MLOs extend from each breadth area in **Figure 3**, we avoid forcing the creation of unhelpful and potentially confused classes just to satisfy the constraints.

It should be no surprise that CCO satisfies the criteria. The eleven ontologies comprising the CCO suite are disjoint ontology hubs - satisfying **HUB** - adopts BFO as a TLO - satisfying **EXTEND** – and extends each of BFO's breadth areas, satisfying **INHERITANCE**. By these criteria, the eleven ontologies that compose the CCO suite collectively count as an MLO.

4. Conclusion

In this paper, we have defended the theoretical and practical underpinnings of BFO and CCO, illustrating how their complementary design principles foster semantic interoperability and improved data quality. By clarifying the distinctions between TLOs, MLOs, and domain extensions, we provided a roadmap for scalable and sustainable ontology engineering. The modular design of CCO, its hub-and-spoke architecture, and its grounding in BFO's realism ensure that domain ontologies can be integrated seamlessly while avoiding common pitfalls like scope creep and ontology silos. The real-world example of aircraft specifications demonstrates the robustness and versatility of the BFO-CCO framework. However, sustaining the success of this ecosystem requires continued methodological vigilance, collaboration among stakeholders, and an emphasis on aligning data quality with semantic rigor. Future work will focus on refining guidelines for MLO development, enhancing support for dynamic data ecosystems, and fostering community adoption of best practices in ontology engineering.

References

[1] Doumanas, D., Soularidis, A., Spiliotopoulos, D., Vassilakis, C., & Kotis, K. (2025). Fine-Tuning Large Language Models for Ontology Engineering: A Comparative Analysis of GPT-4 and Mistral. *Applied Sciences*, 15(4), 2146. https://doi.org/10.3390/app15042146MDPI

[2] Hitzler P, Janowicz K, Tudorache T. Ontology engineering: Current state, challenges, and future directions. *Semantic Web*. 2019;11(1):125-138. doi:10.3233/SW-190382

[3] Jensen, M., Colle, G.D., Kindya, S., More, C., Cox, A.P., & Beverley, J. (2024). The Common Core Ontologies. *ArXiv, abs/2404.17758*.

[4] P. Hitzler, A Review of the Semantic Web Field, Communications of the ACM, 64(2), 76–83, 2021.

[5] The Gene Ontology Consortium, The Gene Ontology Resource: 20 Years and Still Going, Nucleic Acids Research, 47, D330–D338, 2019.

[6] D. Ferrucci et. al., Building Watson: An Overview of the DeepQA Project, AI Magazine, 31(3), 59–79, 2010.

[7] Q. Cao et. al., Ontologies for Manufacturing Process Modeling: A Survey, in: D. Dao et. al. (Eds.), Sustainable Design and Manufacturing 2018, KES-SDM '18, SIST, vol. 130, Springer, pp. 61–70.

[8] E. M. Sanfilippo et. al., Modeling Manufacturing Resources: An Ontological Approach, Product Lifecycle Management to Support Industry 4.0, 540, Springer, 2018 pp. 304–13.

[9] R. Arp, et. al., Building Ontologies with Basic Formal Ontology, MIT Press, 2015.

[10] Smith, B. et al. The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration. Nat Biotechnol 25, 1251–1255 (2007).

[11] J. N. Otte et. al., BFO: Basic Formal Ontology, Applied Ontology, 17(1), 17–43, 2022.

[12] ISO/IEC 21838-2:2021. "Information Technology — Top-Level Ontologies (TLO) — Part 2: Basic Formal Ontology." Accessed Feb 19, 2024.

[13] CUBRC. white Paper—An Overview of the Common Core Ontologies, 2019.
[14] Rudnicki, R., Smith, B., Malyuta, T., & Mandrick, W. Best Practices of Ontology Development. CUBRC, Inc. nist.gov. (2016).
[15] Donohue, B., Jensen, M., Cox, A.P., & Rudnicki, R. (2018). A common core-based cyber ontology in support of cross-domain situational awareness. *Defense + Security*.
[16] Cox AP, Nebelecky CK, Rudnicki R, Tagliaferri WA, Crassidis JL, Smith B. The Space Domain Ontologies.
[17] Moten R. A Practical Approach to Data Modeling using CCO (Common Core Ontologies).
[18] Lori Wade and Craig Martell. Baseline Standards for Formal Ontology within the Department of Defense *and the Intelligence Community*. Chief digital and artificial intelligence officer council members intelligence community chief data officer council members, 25 Jan. 2024
[19] NIST Advanced Manufacturing Series 100-26, 2019.
[20] Keet, Maria C. "Introduction to Ontology Engineering", College Publications, 2020.
[21] T. R. Gruber. A Translation Approach to Portable Ontologies. Knowledge Acquisition, 5(2):199–220, 1993.8.
[22] T. R. Gruber. Toward Principles for the Design of Ontologies Used for Knowledge Sharing. International Journal of Human Computer Studies, 43(5–6):907–928, 1995.
[23] M. R. Genesereth and N. J. Nilsson. Logical Foundations of Artificial Intelli-gence. Morgan Kaufmann, Los Altos, CA, 1987.
[24] Guarino, N., Oberle, D., Staab, S. (2009). What Is an Ontology?. In: Staab, S., Studer, R. (eds) Handbook on Ontologies. International Handbooks on Information Systems. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-92673-3_0
[25] Horgan, J. (n.d.). *Species concepts*. Scientific American. Retrieved May 31, 2025, from https://www.scientificamerican.com/blog/evo-eco-lab/species-concepts/
[26] Mahmoudabadi G, Philops R. (2018). *A Comprehensive and Quantitative Exploration of Thousands of Viral Genomes*. eLife. doi: 1 0.7554/eLife.31955.
[27] Kuhn J. (2020). *Virus Taxonomy*. Reference Modules in Life Sciences. doi: 1 0.1016/B978-0-12-809633-8.21231- 4 .
[28] Baltimore D. (1971). *Expression of Animal Virus Genomes*. Bacteriological Reviews. 35, 235-41.
[29] Brown, Michael E. (2010). How I Killed Pluto and Why It Had It Coming. ISBN 0-385-53108-7.
[30] Clyde W Tombaugh. The search for the ninth planet, Pluto. Leaflet of the Astronomical Society of the Pacific, 5:73, 1946.
[31] T. Bittner, et. al., The Logic of Systems of Granular Partitions, IFOMIS, 2005.
[32] Correia F. Ontological dependence. *Philosophy Compass* 3.5. 2008. p. 1013-1032.
[33] McKitrick J. A case for extrinsic dispositions. *Australasian Journal of Philosophy*. 2003.
[34] Doran P, Tamma V, Iannone L. Ontology Module Extraction for Ontology Reuse: An Ontology Engineering Perspective. 16th ACM International Conference on Information and Knowledge Management; 2007. pp. 61–70.

[35] Orbst, L. et. al. (2010) Introduction: Ontologies, Semantic Technologies, and Intelligence. Ontologies and Semantic Technologies for Intelligence.
[36] Smith, Barry; Ceusters, Werner. (2015). "Aboutness: Towards Foundations for the Information Artifact Ontology". In *Proceedings of the Sixth International Conference on Biomedical Ontology (ICBO).*
[37] Randell, D. A., Cui, Z., & Cohn, A. G. (1992). A spatial logic based on regions and connection. In B. Nebel, C. Rich, & W. Swartout (Eds.), *Proceedings of the 3rd International Conference on Knowledge Representation and Reasoning (KR'92)* (pp. 165–176). Morgan Kaufmann.
[38] Chris Mungall, David Osumi-Sutherland, James A. Overton, Jim Balhoff, Clare72, pgaudet, Matthew Brush, Nico Matentzoglu, Vasundra Touré, Damion Dooley, Michael Sinclair, Anthony Bretaudeau, Scott Cain, Melissa Haendel, diatomsRcool, Jen Hammock, Marie-Angélique Laporte, Mark Jensen, & Martin Larralde. (2020). oborel/obo-relations: 2020-07-21 (v2020-07-21). Zenodo. https://doi.org/10.5281/zenodo.3955125
[39] Beverley, J., Colle, G.D., Jensen, M., Benson, C., & Smith, B. (2024). Middle Architecture Criteria. *ArXiv, abs/2404.17757*.
[40] Comments for INCITS SubGroup on MLO CCO Wiki. (2020, December 18). GitLab. https://opensource.ieee.org/cco/CommonCoreOntologies/-/wikis/Comments-for-INCITS-SubGroup-on-MLO-CCO
[41] Drobnjakovic, M., et. al., The Industrial Ontologies Foundry (IOF) Core Ontology, Formal Ontologies Meet Industry (FOMI) 2022.
[42] De Colle, Giacomo, et. al. Ontology Development Strategies and the Infectious Disease Ontology Ecosystem. Proceedings of the International Conference on Biomedical Ontologies 2023.
[43] Immaterial entity. 2012. Issue #21. BFO-ontology/BFO. GitHub. https://github.com/BFO-ontology/BFO/issues/21