

IDETC/CIE2023-116866

RECENT DEVELOPMENTS IN ONTOLOGY STANDARDS AND THEIR APPLICABILITY TO BIOMANUFACTURING

Milos Drobnyakovic, Boonserm Kulvatunyou, Simon Frechette, Vijay Srinivasan

National Institute of Standards and Technology, Gaithersburg, Maryland, USA

ABSTRACT

ISO and IEC have jointly initiated, and recently issued, a series of standards (the ISO/IEC 21838 series) for top-level ontologies. These standards have been used by industrial consortia to develop and disseminate standards for mid-level ontologies to ease the development of ontologies in specific domains and applications. This paper traces the motivation for these concerted efforts to the successful application of ontologies in biological sciences. It then analyzes these recently released international standards for top-level and mid-level ontologies, and shows how they can be used for domain-level and application-level ontologies. Finally, it provides a rationale for applying these standards to the important and emerging field of biomanufacturing.

Keywords: ISO, IEC, OBO Foundry, BFO, CCO, IOF, standards, ontology, biology, biomanufacturing

1. INTRODUCTION

It is a general practice in the development of a standard to start with some definition of terms used in the domain covered by the standard. These terms and their definitions are discussed and voted on by subject matter experts in that domain of discourse and published as a documentary standard by an appropriate agency (e.g., ISO 23247 series of standards for digital twins) [1]. Definition of the terms specific to the standard requires the use of other terms, whose meanings are assumed to be either well-known in common usage or defined in other accepted standards known to the experts. It is not surprising that the subject matter experts have the final say in the domain covered by the standard. The terms thus defined form the basis for developing schemas using data modeling languages, such as XML Schema, for data exchange among computerized information systems (e.g., AgGateway Open Standards and connectSpec) [2,3,4].

While this time-honored practice of a decentralized and asynchronous approach to the definition of terms in standards

is still very much in vogue, some recent events have taken more rigorous and formal steps towards terminology development using ontologies – in short, to put ‘terminology on steroids.’ Broadly speaking, ontology starts with terms and their definitions within a given domain and expands them to a body of knowledge about that domain. The field of ontology has a long history in philosophy and comes with considerable baggage. More recently, however, advances in information science and computations have made some inroads into how ontologies can be made more computable [5]. This has opened the possibilities for greater rigor and automation in the way terms are defined and shared across multiple standards and multiple domains. Since data are strongly associated with terms, this has also opened the path to greater integration and interoperability of information systems used in various scientific and industrial applications.

Such potential gains, however, require much greater upfront effort to organize the approach to define terms. One solution to this complex problem is to adopt a tiered architecture, where a top-level ontology will tackle a few most commonly used terms (also referred to as ‘highly general terms’) [6, 7]. These terms can then be progressively specialized hierarchically in lower tiers (in a process known as the ‘hub-and-spoke’ approach) for additional terms till an application-specific standard is reached. An important requirement in this approach is that the *relationship* between the terms should also be clearly defined. Benefits can be demonstrated only when an application-specific standard (for example, in a field of manufacturing) can use this tiered approach to achieve its objectives faster, cheaper, or better than what could have been achieved by current conventional means.

This paper provides an analysis of recent standards that define the top-level and mid-level ontologies and discusses their applicability to the emerging biomanufacturing industry. Targeting biomanufacturing may appear to be a curious choice, but it turns out that some of the initial successes with ontologies over the past two decades have been in their application to

biosciences, for example using the gene ontology [8]. It is thus a worthwhile endeavor to examine how much of the initial success of applying ontology in biosciences can be duplicated in biomanufacturing.

The rest of the paper is organized as follows. Section 2 provides a brief history of the successful application of ontologies in biological sciences. In Section 3, the recently released international ontology standards are analyzed to illustrate their hub-and-spoke approach. Section 4 addresses the applicability of these ontology standards to the important and emerging biomanufacturing field. Some promising near-term developments are proposed in Section 5. Specifically, we outline the potential to apply ontologies to data driven modeling, digital twins, and circular bioeconomy. Finally, a summary and concluding remarks are given in Section 6.

2. A BRIEF HISTORY OF ONTOLOGY IN BIOLOGY

Biological organisms are the most complex adaptive systems known to humans. In 1994 Murray Gell-Mann, a Nobel Laureate in particle physics, exclaimed “Even the jaguar, the ancient symbol of power and ferocity, is a bundle of quarks and electrons, but what a bundle! It exhibits an enormous amount of complexity, the result of billions of years of biological evolution” [9]. In 1997 Marin Rees, the United Kingdom’s Astronomer Royal, quipped “It is more difficult to understand a frog than the cosmos” because the biology of frog has several layers of structure [10].

This formidable complexity of biology has been tamed to some extent by modern developments in the last few decades. Thanks to breakthroughs in modern biology, the world has a much better mechanistic and informational understanding of biological systems. Advances in molecular biology and genomics have revealed the inner workings of living cells and genes to an extent that a living cell now looks like a sophisticated factory that runs on genetic information [11]. Inside the cells, complex transcription and translation machinery (composed of biological macromolecules such as enzymes and rRNA) act to manufacture large biomolecules called proteins based on the information encoded in the genome.

It was during the process of unraveling the genetic and cellular complexity that a need arose to capture and understand vast amounts of data coming out of biological research. DNA sequencing had by then revealed that a large fraction of the genes is shared by different organisms such as yeasts, fruit-flies, mice, and humans. Since proteins are encoded by genes and proteins perform most of the biological functions, it was strongly felt that the knowledge of biological roles of such shared proteins in one organism can be transferred to other organisms. This prospect was, however, hampered by the fact that experts working on different biological organisms used different terms for the genes and their biological functions. The information systems created by these domain experts could not interoperate because of this terminological difference.

An interesting choice was made at this stage by the bio-research community to attack this information conundrum.

A group of researchers well versed in bioinformatics decided to apply advances in ontology to tackle the terminology divergence problem. In 1998 they formed a Gene Ontology (GO) Consortium, which undertook the task of producing a dynamic, controlled vocabulary that could be applied across multiple biological organisms even as the knowledge of genes and proteins and their roles in cells was still accumulating and changing. Towards the end of the 20th century, the GO consortium produced the following three ontologies:

1. **Molecular function**: The terms in this ontology describe activities that occur at the molecular level. For example, these can be ‘catalysis’ and ‘transport.’
2. **Cellular component**: The terms in this ontology refer to the location in the cellular structure in which a gene product (e.g., a protein) performs a function. For example, this can be ‘mitochondrion.’
3. **Biological process**: The terms in this ontology refer to processes such as ‘DNA repair.’

Since the beginning of the 21st century, GO has evolved beyond being merely a vocabulary of terms and definitions [5]. Two major developments in this evolution are:

1. **Classes and relations**: The ‘terms’ are now called ‘classes’ and the *relations* between the classes (terms) have been given greater attention. Relations currently used in GO are:
 - i. *is a*: This relation forms the basic structure of GO. If *A is a B*, then node *A* is a subtype of node *B* (that is *A* logically subsumes *B*, in description logic parlance).
 - ii. *part of*: This relation is used to represent part-whole relationship.
 - iii. *has part*: This relation is the logical inverse of the *part of* relation.
 - iv. *regulates*: This relation describes the case in which one process directly affects the manifestation of another process or quality, that is, the former regulates the latter.

A directed graph can be created with classes as the nodes and relations between classes as the arcs. Fig. 1 illustrates a few basic biological terms using a very simple GO graph, where the mitochondrion is an organelle and it is also a part of cytoplasm. Most GO graphs tend to be more complex, consisting of hundreds of nodes and arcs. The GO graph structure immediately suggests a graph database, and an information modeling system using the ‘subject-predicate-object’ model.

2. **Annotations**: These are detailed explanations about the function of a gene – such as how a gene functions at the molecular level, where in the cellular component it functions, and what biological processes it helps carry out. In addition, the annotation statements contain references, evidence, and any other supporting document. Taken

together, these statements provide a snapshot of current biological knowledge about genes.

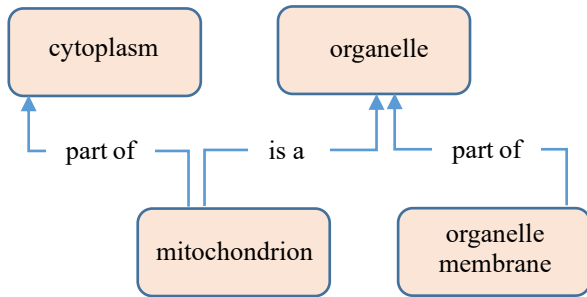


FIGURE 1. A SIMPLE GENE ONTOLOGY GRAPH [5].

An even more important contribution of GO is that it spawned other active ontological research, such as Open Biological and Biomedical Ontologies (OBO) for several more ontologies (e.g., cell ontology, protein ontology, vaccine ontology) and the OBO Foundry (since 2007) to impose some discipline on how these bio-related ontologies are created and managed [12]. In fact, GO itself has reformed and is now a member of the OBO Foundry. Drawing from the experience of managing several bio-related ontologies, the OBO Foundry has established the following overall principles (with far-reaching consequences for ontology standards outlined in Section 3):

- **Open:** The ontology must be openly available to be used by all without any constraint other than (a) its origin must be acknowledged and (b) it is not to be altered and subsequently redistributed in an altered form under the original name or with the same identifiers.
- **Common Format:** The ontology is made available in a common formal language in an accepted concrete syntax.
- **URI/Identifier Space:** Each ontology must have a unique IRI in the form of an OBO Foundry permanent URL (PURL).
- **Versioning:** The ontology provider has documented procedures for versioning the ontology, and different versions of ontology are marked, stored, and officially released.
- **Scope:** The scope of an ontology is the extent of the domain or subject matter it intends to cover. The ontology must have a clearly specified scope, and contents that adhere to that scope.
- **Textual Definitions:** The ontology has textual definitions for the majority of its classes and for top level terms in particular.
- **Relations:** Relations should be reused from the Relations Ontology (RO).
- **Documentation:** The owners of the ontology should strive to provide as much documentation as possible.

- **Documented Plurality of Users:** The ontology developers should document that the ontology is used by multiple independent people or organizations.
- **Commitment To Collaboration:** OBO Foundry ontology development, in common with many other standards-oriented scientific activities, should be carried out in a collaborative fashion.
- **Locus of Authority:** There should be a person who is responsible for communications between the community and the ontology developers, for communicating with the Foundry on all Foundry-related matters, for mediating discussions involving maintenance in the light of scientific advance, and for ensuring that all user feedback is addressed.
- **Naming Conventions:** The names (primary labels) for elements (classes, properties, etc.) in an ontology must be intelligible to scientists and amenable to natural language processing. Primary labels should be unique among OBO Library ontologies.
- **Maintenance:** The ontology needs to reflect changes in scientific consensus to remain accurate over time.

During the process of implementing these principles, the OBO Foundry found that it was advantageous to adopt a Basic Formal Ontology (BFO) that can stand at the top of all ontologies [13]. This was an important insight because it pointed towards a tiered architecture for developing ontology standards in various domains (as outlined in Section 3).

It turns out that BFO can be small and domain-neutral (that is, not specific to biology at all). It has a relatively small number of classes (terms) and relational expressions (relations) when compared with domain-specific ontologies. The BFO classes define terms such as ‘entity’, ‘quality’, ‘process’, ‘spatial region’, and ‘temporal region’; these terms pertain to all domains of reality. Examples of BFO relations include ‘located in’, ‘participates in’, and ‘exists at’; these relations allow for assertions about these realistic classes. In addition, BFO is formalized in OWL 2 (Web Ontology Language 2) and CL (Common Logic) to enable automated reasoning, consistency checking, mapping/translating to different views and formats, and machine learning. These developments have been coordinated with recent efforts in ISO/IEC standardization, which is addressed in the next section.

3. ANALYSIS OF RECENT ONTOLOGY STANDARDS

The emergence of domain-neutral ontologies, such as BFO, prompted a Joint Technical Committee of ISO (International Organization for Standardization) and IEC (International Electrotechnical Committee) to start developing and issuing a series of standards on Top-Level Ontologies (TLO), the first two of which were released in 2021 [6, 7]. In parallel to these developments, two activities were undertaken to standardize and release mid-level ontologies (MLO) that build on the TLO standards. Using a tiered architecture, TLO and MLO can then be used to build domain-dependent ontology standards for

various applications. The levels in the tiered architecture of these ontology standards are described in the following subsections.

3.1 Top-Level Ontologies

Two major tasks drive the standardization of top-level ontologies (TLO). Firstly, as mentioned earlier, just a few terms (classes) and relations (relational expressions) were found to be common among various bio-related ontologies and these were also found to be independent of the domain of discourse; these terms and relations should be collected and agreed upon. Secondly, while ontologies may very well serve the purpose of providing a snapshot of the current knowledge in any particular domain, the need for exchanging data among various information systems used in science and industry demands greater mathematical formalism and computational support for these terms and relations.

The ISO/IEC 21838 series of standards address these two tasks by starting with a set of requirements in Part 1 [6]. These requirements stipulate that a TLO standard must provide:

- TLO as a textual artefact. This means that a TLO has a natural language documentation that provides (1) a list of domain-neutral terms and relational expressions, incorporating identification of primitive terms, and (2) definitions of the meanings of the terms and relational expressions listed.
- Axiomatization in the Web Ontology Language (OWL 2). This means that a TLO is made available via at least one machine-readable axiomatization in OWL 2 with the direct semantics, or in some description logic that is designated by W3C as a successor of OWL 2.
- Axiomatization in a language conforming to Common Logic (CL). This means that the TLO is made available via an axiomatization in a language conforming to ISO/IEC 24707 [14].

Requiring axiomatization in both CL and OWL 2 may seem superfluous, but they serve different purposes:

- CL is a logical framework with the full expressivity of first-order logic (FOL), the unifying framework for all semantic web applications. Formalization in a language with the expressivity of FOL is required because weaker expressivity would not allow the ontology to capture in a formal way the implications of axioms in areas such as mereology (that is, part-whole relationship) and theories of location and change.
- Formalization in a language like OWL 2, even though it is less expressive than CL, is needed because it is decidable. This means that it can be used effectively by software systems for logical reasoning and ontology quality assurance. One of the requirements of ISO/IEC 21838-1 is

that documentation must be provided specifying how it is shown that the OWL 2 axiomatization is logically derivable from the CL axiomatization.

In addition, ISO/IEC 21838-1 stipulates several documentation practices similar to those found in the OBO Foundry principles described in Section 2.

There are a few TLOs that have already found favor with software applications developers. One such TLO is the Basic Formal Ontology (BFO) mentioned in Section 2. BFO has been picked up by Part 2 of ISO/IEC 21838 [7] for standardization. Since Part 2 should conform to the requirements set forth in Part 1, further enhancements to BFO were undertaken and tested for conformance to ISO/IEC 21838-1. The entire body of BFO that is covered by ISO/IEC 21838-2 is now available for open access at an ISO website [15], which includes:

- An Excel spreadsheet that contains the textual description of BFO. Currently, it has 84 terms and relational expressions, confirming the relative compactness of this TLO. Some of the terms and relational expressions in the ISO/IEC BFO spreadsheet are shown in Tables 1 and 2, respectively, along with explanation/definition and examples. Note that only a few of the 84 terms and relational expressions are shown in Tables 1 and 2.
- OWL 2 files and CL files that contain axiomatized versions of BFO.

ISO and IEC are now working on additional parts of the ISO/IEC 21838 series to standardize other TLOs. At present, Part 3 on DOLCE (Descriptive Ontology for Linguistic and Cognitive Engineering) and Part 4 on TUpper are under development.

Some of the terms and their definitions in Table 1 may not be familiar, or may be puzzling, to people outside the ontology community. ‘Entity’ is a term that is used quite extensively in standards, especially in those that would like to apply information technology. Yet, its definition in Table 1 as ‘anything that exists or has existed or will exist’ may seem rather tame. But it is the best domain-neutral definition that could be found.

Other terms in Table 1 that may be unfamiliar to outsiders are ‘continuant’ and ‘occurrent.’ These terms play important roles in dividing the ‘entity’ into two disjoint domains as can be seen from the definitions of the rest of the terms in Table 1. The use of the terms ‘continuant’ and ‘occurrent’ with any other term has the consequence of imposing a time-related clarification of that term. The use of ‘continuant’ and ‘occurrent’ continues into the definition of relational expressions as well, as can be seen from Table 2 in the relational expressions ‘continuant part of’ and ‘occurrent part of.’ A quick perusal of the 84 terms and relational expressions of ISO/IEC 21838-2 archived in [12] reveals that the concepts of continuant and occurrent are quite pervasive.

It was already remarked that the terms and relational expressions of TLO must be axiomatized in OWL 2 and CL. As an example of such axiomatization, consider the transitivity of

the relational expression ‘located in’ in Table 2. Using mathematical notations, an axiom for this will look like this:

Axiom: ‘located in’ is transitive at a time:

$$\forall a, b, c, t, t_2 (\text{locatedIn}(a, b, t) \wedge \text{locatedIn}(b, c, t_2) \wedge \text{temporalPartOf}(t, t_2) \rightarrow \text{locatedIn}(a, c, t))$$

TABLE 1. SOME OF THE TERMS STANDARDIZED IN ISO/IEC BFO [12].

Term	Explanation/Definition	Examples
entity	An entity is anything that exists or has existed or will exist	Julius Caesar, the Second World War, your body mass index, Verdi’s Requiem.
continuant	A continuant is an entity that persists, endures, or continues to exist through time while maintaining its identity	A human being, a tennis ball, a cave, a region of space, someone’s temperature.
occurrent	An occurrent is an entity that unfolds itself in time or it is the start or end of such an entity or it is a temporal or spatiotemporal region	The life of an organism from the beginning to the end of its existence.
spatial region	A spatial region is a continuant entity that is a continuant part of the spatial projection of a portion of spacetime at a given time	A cube-shaped region of space, a sphere-shaped region of space, the region of space occupied by all and only the planets in the solar system at some point in time.
temporal region	A temporal region is an occurrent over which processes can unfold	The time during which a cell-division occurs.
material entity	A material entity is an independent continuant that at all times at which it exists has some portion of matter as continuant part	A human being, the undetached arm of a human being, an aggregate of human beings.
process	p is a process means: p is an occurrent that has some temporal proper part and for some time t, p has some material entity as participant at t	An act of selling, the life of an organism, a process of sleeping, a process of cell-division, a beating of the heart, a process of meiosis, the taxiing of an aircraft, the programming of a computer.
quality	A quality is a specifically dependent continuant that, in contrast to roles and dispositions, does not require any further process in order to be realized	The colour of a tomato, the ambient temperature of this portion of air, the length of the circumference of your waist, the shape of your nose, the shape of your nostril, the mass of this piece of gold.

TABLE 2. SOME OF THE RELATIONAL EXPRESSIONS STANDARDIZED IN ISO/IEC BFO [12].

Relational Expression	Explanation/Definition	Examples
is a	A is a B =Def. for all x, t, if x is an instance of A at t then x is an instance of B at t	House is a building, symphony is a musical work of art, promenade is a dance step, promise is a speech act, cell division is a biological process.
continuant part of	b continuant part of c at t means: b and c are continuants & b is a part of c at t	Mary's arm is a continuant part of Mary in the time of her life prior to her operation; the Northern hemisphere of the planet Earth is a continuant part of the planet Earth at all times at which the planet Earth exists.
occurrent part of	b occurrent part of c =Def. b is a part of c & b and c are occurrents	Mary's 5th birthday is an occurrent part of Mary's life; the first set of the tennis match is an occurrent part of the tennis match.
located in	b located in c at t =Def. b and c are independent continuants and not spatial regions, and the spatial region which b occupies at t is a (proper or improper) continuant part of the spatial region which c occupies at t	Your arm is located in your body; this stem cell is located in this portion of bone marrow; this portion of cocaine is located in this portion of blood; Mary is located in Salzburg; the Empire State Building is located in New York.
exists at	exists at is a relation between a particular and some temporal region at which the particular exists	First World War exists at 1914-1916, Mexico exists at January 1, 2000.

In simple words, the transitivity axiom states that if a is located in b and b is located in c , then a is located in c . To be rigorous, however, one has to introduce a notion of time. For this, one can use the relational expression ‘temporal part of’ defined in the ISO/IEC BFO. Then, $\text{temporalPartOf}(t, t2)$ in the Axiom above means, in simple words, that the time or time period t is within (that is, part of) the time or time period $t2$. So, the transitivity axiom states that:

For all (applicable) $a, b, c, t, t2$,
 if a is located in b at time t , and
 b is located in c at time $t2$, and
 t is within $t2$,
 then a is located in c at time t .

The transitivity is now much clearer and more rigorous. What may at first appear to be a cumbersome way of stating something simple turns out to be the best and unambiguous way to state it. Axiomatization in CL and OWL 2 provide computable representations of such axioms for BFO, which can then be processed automatically by software systems for applications such as consistency checking, automated reasoning, and machine learning.

3.2 Mid-Level Ontologies

The path from a Top-Level Ontology to an application-level ontology, such as a biomanufacturing ontology, is eased by a Mid-Level Ontology (MLO) that uses TLO to build more terms and relational expressions that get closer to the application. This was the motivation behind two of the MLOs that have been developed in recent years. Both depend on BFO as the TLO, and are described below:

- Common Core Ontologies (CCO) [16]: As the name implies, CCO is not a single ontology but is comprised of the following 11 distinct open-source ontologies that have varying degrees of interdependency: Information Entity Ontology, Agent Ontology, Quality Ontology, Event Ontology, Artifact Ontology, Time Ontology, Geospatial Ontology, Units of Measure Ontology, Currency Unit Ontology, Extended Relation Ontology, and Modal Relation Ontology. In total, across the 11 ontologies, there are 1394 classes (that is, terms) and 225 relational expressions, with the axiomatization provided only in OWL 2. Within this wide array of classes, several are vital to the domain of manufacturing (e.g., agent, state, material artifact, artifact design, prescribes, and plan). As such, CCO has been used as a basis of a few domain-specific manufacturing ontologies as well as ‘proof of concept’ papers. However, CCO aims to cover a much broader territory than manufacturing and thus leaves open some of the other terms and relations critical to manufacturing.
- Industrial Ontologies Foundry Core (IOF Core) [17]: Industrial Ontologies Foundry (IOF) was an effort initiated

in 2016 with the help of NIST (National Institute of Standards and Technology, USA) to address the needs of the industrial sector, especially the manufacturing industry. Inspired by the model of OBO Foundry and driven by the industrial need to connect data among industrial information systems, IOF started working on an MLO by convening workshops and teaming up with standards development organizations such as the Open Applications Group Inc (OAGi) [18].

The MLO that resulted from these efforts was released as IOF Core Ontology Version 1 beta in 2022 [17]. IOF Core, as it is called, currently has 57 terms and 38 relational expressions, with the goal of axiomatization in OWL 2 and CL. For now, the IOF Core pursues an ‘OWL first’ policy, with the result that the CL is not used to its full extent (e.g., creating non-binary relations). Formalizations are currently constructed in a manner that aims to further clear up and elaborate the OWL 2 axioms and not introduce a complete complementary system in CL as is currently given in BFO.

Table 3 shows some of the terms defined in the IOF Core. In Table 4, some of the relational expressions defined in the IOF Core are explained. The terms and relational expressions standardized by the IOF Core have been chosen after considerable discussion with industrial experts and the IOF Core will continue to evolve in future versions.

3.3 Domain- and Application-Level Ontologies

Domain-level ontologies address a specific domain of interest (e.g., supply chain, production planning, biological cell line development). Application-level ontologies stem from domain-level ontologies, are even narrower in scope, and are typically focused on one use-case (e.g., supply chain for semiconductor chips used in smartphones, production planning for an electric vehicle that uses lithium-ion batteries, a CAR-T therapy cell line).

Consequently, domain- and application-level ontologies introduce very specific and narrow concepts that permit higher ontological commitment of terms and relational expressions (ideally formalized as axioms) than mid-level or top-level ontologies. This high ontological commitment on interpretation enables precise definition of terms and relational expressions, and thus allows a significant increase in consistency checking and uniformity of usage.

However, if domain and application ontologies are developed as stand-alone ontologies, there is a danger of over-constraining the interpretation of some of the more general terms these ontologies utilize. Consequently, their interoperability with other ontologies that deal with a related area can be significantly hampered, limiting their overall utilization. This is the prime motivator to base them hierarchically, using a hub-and-spoke approach, on TLO and MLO.

TABLE 3. SOME OF THE TERMS STANDARDIZED IN IOF CORE [14]

Term	Definition	Examples
design specification	information content entity that prescribes something human-made	document specifying the characteristics of a pharmaceutical product; the design of a software program to schedule the work orders in a factory
material artifact	object that is deliberately created to have a certain function	a machine; a screwdriver; a screw; a sheet of paper
engineered system	system that is deliberately created to have certain function	machine, laptop, traffic light system
equipment role	role held by a material artifact when it is planned to be involved in or is involved in carrying out some part of a planned process and that is not consumed in that planned process	role of a wrench when it is used in the maintenance process of a car; role of a chromatography column that is planned to be used in a protein purification process; role of a truck that is used to transport goods to the buyer; role of a single-use bioreactor when it is planned to be used in the upstream phase of a biomanufacturing process
product production process	business process that consists of at least one manufacturing process through which raw materials and components are transformed or modified to create a material product	making of an engine block as a product that consists of many processes such as machining process, assembly process, inspection process etc.

TABLE 4. SOME OF THE RELATIONAL EXPRESSIONS STANDARDIZED IN IOF CORE [14]

Relational Expression	Definition	Examples
acts on behalf of at some time	relation from a material entity to a person or a group of agents or engineered system that holds when the material entity participates in some planned process in order to fulfill an objective for the person or group of agents or engineered system	an agent acts on behalf of a business organization; a lawyer acts on behalf of a person.
satisfies requirement	relation from an entity to a requirement specification that the entity conforms to	a piece of software satisfies a UML requirement specification, a design specification of a car satisfies its functional requirement specification
is available to at some time	relation from a material entity or physical location to an agent that holds when the material entity or physical location have a capability that is needed by the agent to fulfill some objective carried by the agent	a roll of aluminum (resource) is available to an agent to use in a forming process, a milling workstation (resource) is available to a manufacturer to produce some parts
has output	relation from a process to someone or something physical or digital (continuant) that participates in the process such that it is generated or modified during the process, and such that it exists at the end of the process	chemical manufacturing process has output a waste stream; toluene manufacturing process has output a certain quantity of toluene; car manufacturing process has output a car

Several domain- and application-level ontologies have been built using this hierarchical hub-and-spoke approach. For example, a Supply Chain Reference Ontology (SCRO) [19], which can be viewed as a domain-level ontology, is based on the IOF Core. It defines the terms and relationships for the structure (members and their roles, functions, capabilities, relations, and resources) and operations (processes and flow of material and information) of supply chains. Stemming from SCRO, an application-level ontology titled Supply Chain Traceability (SCT) ontology has been created [20]. SCT is an ontology that deals explicitly with supply chain traceability within the agricultural supply chain.

Another IOF Core based ontology is a Reference Ontology for Maintenance Work Management [21]. An additive manufacturing ontology (AMO) is an example of a CCO-based ontology which deals with the lifecycle of additive manufacturing products [22].

There are also other manufacturing-related ontologies, such as a Machine-Tool Model Ontology [23] and Ontology for Describing Production Equipment [24], that do not, or only partially, adhere to the hub-and-spokes approach. Hence their interoperability and utility outside the intended use cases seem to be limited.

4. APPLICABILITY OF ONTOLOGY STANDARDS TO BIOMANUFACTURING

Broadly speaking, the term bioeconomy refers to the share of the economy based on products, services, and processes derived from biological resources (such as micro-organisms). It is predicted that bioeconomy will be a big component of the future global economy [25]. Therefore, it is not surprising that many countries and multi-national organizations have defined the term bioeconomy in more detail to drive policy decisions.

One such definition, as defined by the National Academies of Science, Engineering, and Medicine (NASEM), states that bioeconomy is “economic activity that is driven by research and innovation in the life sciences and biotechnology, and that is enabled by technological advances in engineering and in computing and information sciences.” Specifically, the U.S. bioeconomy is built on three biomanufacturing sectors: *bioindustrial*, *biomedical*, and *agricultural* [25].

Since bioeconomy and biomanufacturing are so closely aligned with life sciences and biotechnology, it is natural to expect that much of the ontological developments described in Section 2 to support bio-related research will have some impact on biomanufacturing. In addition, the ontological developments to support the manufacturing industry described in Section 3 can also be expected to play an important role in biomanufacturing. Most of the manufacturing-related terms and relational expressions critical to biomanufacturing are not available in the BFO (described in Section 3.1) and can only be found in an MLO such as the IOF Core (described in Section 3.2).

It is in this context that several data-related problems that plague all the three sectors of biomanufacturing were identified. One of the primary obstacles faced by data utilization in

biomanufacturing is the lack of connectivity between data associated with terms and their subsequent transformation into knowledge and insights. Specifically, interconnection of techno-economic data that supports analysis of cost, performance, sustainability, and production metrics is currently lacking. This lack of interconnection can be attributed to the following challenges encountered while utilizing data associated with biomanufacturing:

1. Data is currently stored in many different formats that are not necessarily mutually coherent.
2. There is a lack of explicit connections between multivariate statistics model results and the ‘physical and biochemical causality.’
3. High complexity and volume that stems from ‘omics’ (such as genomic and proteomic) data.
4. Lack of metadata in many data sources.

As such, any subsequent data utilization in understanding productivity impact or sustainability potential is incredibly laborious and, in many cases, has to be tailored for each process. The four generic data challenges outlined above are common across all the three sectors of biomanufacturing and will be referred to again by their numbers in the following subsections.

In addition, each biomanufacturing sector has its own distinct set of data challenges that stem from its specific objectives and regulatory environment. The rest of this section outlines each sector's key data challenges and analyzes how different ontologies could be used to alleviate the associated pain.

4.1. Applying Ontologies to Bioindustrial Manufacturing

Bioindustrial manufacturing can be defined as the manufacturing of chemicals and/or materials through the utilization of biosynthetic or semi-biosynthetic routes. Historically, bioindustrial manufacturing has been primarily focused on products that cannot be produced by traditional chemical routes (e.g., wine), or those products that have simple structures and are already byproducts of naturally existing microorganisms (e.g., ethanol).

However, with rapid advancements in genetic engineering and next-generation sequencing approaches, bioindustry has started to create manufacturing routes for more complex products (e.g., fine chemicals and high-value precursors) and novel bio-based materials (e.g., materials for additive manufacturing, bioplastics). In addition to technological advancements, the primary driving force behind the development of new biologically produced products is that there is a high potential to reduce waste and thus make the production greener and more sustainable. Also, unique bio-based products have a higher recyclability potential than traditional materials such as plastics.

One of the obstacles standing in the way of realizing these potentials is the set of data challenges mentioned earlier. Also,

bioindustrial manufacturing would greatly benefit from establishing explicit connections between process configurations (e.g., particular process parameters, genetic modification of a cell line) and economic analysis metrics. The reason is that the bioindustrial sector is competing in productivity and quality with traditional commodity manufacturing. Thus, it needs to ensure the creation of competitive processes for its widespread adoption. As has been described in the previous Section 3, ontologies have the potential to ameliorate most of the data challenges. Nevertheless, literature on ontologies that specifically address bioindustrial manufacturing processes and bioindustrial product lifecycle is relatively scarce and just starting to emerge.

More specifically, most ontologies related to bioindustrial manufacturing address a very narrow part of the entire landscape. For instance, Menke et al. have developed a draft ontology specific to biocatalysis [26]. Gruhn et al. have created an ontology that addresses biocatalysis process scale-up [27]. Caño De Las Heras et al. released a proof-of-concept project [28] of a non-relational database of process models, which rests on an ontology – details of which are currently not available.

The emerging consensus is that an ontology that provides a set of foundational concepts of biomanufacturing would have much better impact. More specifically, the ontology would be able to provide the basis for (1) standardization of concepts contained in data, and (2) explicitly linking such data with a specific production phase as well as with other relevant information, thereby also enriching the metadata (hence effectively helping with data challenges 1, 2, and 4 mentioned earlier). From it, by following the hub-and-spokes approach, specialized ontologies can be established which handle more specific cases. For instance, an ontology that deals with the biomanufacturing of a specific chemical (e.g., styrene) could be created.

Having said that, any potential work in creating a bioindustrial manufacturing ontology does not have to start from ‘ground zero’ and can reuse and extend concepts from related fields. A case in point is the traditional chemical industry. Unlike in the biomanufacturing field, the chemical industry is more ‘ontologically developed.’ An example is Process Chemistry Ontology, which models the domain of process chemistry from route scouting, process optimization, process validation, and process maintenance [29]. Many of these concepts are also relevant and common to all three biomanufacturing sectors.

In addition, portions of ontologies developed in the biomedical research field cover certain aspects of biomanufacturing. For example, Allotrope Ontology has many terms and relations relevant to laboratory analysis and experimentation [30]. Terms and relations within the Cell Line Ontology [31] could potentially be reused in annotating and ontologically representing biomanufacturing cell-lines. Human Proteome Organization (HUPO)-Proteomics Standards Initiative (PSI) has created an ontology for a mass spectrometry analysis pipeline for proteomics (one of the ‘omics’ areas) [32]. This ontology could be aligned with a biomanufacturing

ontology with specific constructs that would provide more connections to proteomics and thus potentially ease data challenge 3 mentioned earlier.

Hence, a bioindustrial manufacturing ontology could be 1) developed in compliance with the already existing literature and 2) in a manner that is aligned with a broader top-level and mid-level ontology. This would facilitate greater interoperability and reuse and, conversely, enable a biomanufacturing ontology to have a firm semantic and formal grounding in many related areas shared across all three biomanufacturing sectors.

4.2. Applying Ontologies to Biomedical Manufacturing

Biomedical manufacturing can be defined as a union of two distinct manufacturing sectors:

1. *Biopharmaceutical manufacturing* is the manufacturing of pharmaceuticals through the utilization of living organisms.
2. *Biomedical device manufacturing* is the manufacturing of medical devices that utilize living organisms or the associated biochemical molecules (e.g., enzymes) as active components.

Biopharmaceutical manufacturing is by far the larger of the two sectors. It encompasses the production of a wide array of products from ‘relatively small and simple’ peptides such as insulin to entire cells such as CAR-T cells. Such products have varying manufacturing requirements, configurations, and degrees of available process knowledge (e.g., the process parameters of antibody manufacturing are much better understood compared to CAR-T Cell manufacturing). Nevertheless, many of the data challenges described in the rest of this subsection are common across all biomedical product types.

The role of data in biopharmaceutical manufacturing and its subsequent utilization is expanding in many directions in recent times. For example, increasing compliance and validation requirements have led to an increase in the amount of documentation that must be filed to prove that the process complies with regulatory standards. Moreover, the biopharmaceutical industry is shifting towards Smart Manufacturing (also known as Industry 4.0) – a manufacturing paradigm shift aiming to minimize human intervention through advanced inter-machine communication, real-time data collection, and automated data-driven decision making.

The data challenges for the biopharmaceutical sector intersect with those for the bioindustrial sector already described in the previous Subsection 4.1. That is, data challenges 1, 2, and 4 also greatly impact the biopharmaceutical industry. One special impact on the future direction is the data challenge 1 as the biopharmaceutical industry is transitioning to continuous manufacturing – a method of production where inter-unit data communication is imperative for success.

Biopharmaceutical manufacturing also has a few other unique challenges. The first is that data from clinical trials

should be considered and used synchronously with process-based data. The second stems from the highly regulated nature of the industry which requires a higher degree of data provenance (that is, data sources and their trustworthiness).

Despite the potential benefits of using an ontological formalism, the current ontological landscape of biopharmaceutical manufacturing is quite similar to that of bioindustrial manufacturing. Specifically, most of the ontological work is focused on the product, with no link to the manufacturing process. The Vaccine Ontology [33], for example, aims to support basic vaccine research and development and clinical vaccine usage but does not provide any further connection to the vaccine manufacturing lifecycle. Ontologies that address biopharmaceutical manufacturing are scarce, developed independently, and have limited domain coverage. As an example, Coakely et al have developed an ontology that covers the determination of CQAs (Critical Quality Attributes) and CPPs (Critical Process Parameters), but is lacking in terms of control system representation, monitoring and sampling strategies, and connections to clinical trial data [34].

Biopharmaceutical manufacturing, in the same vein as bioindustrial manufacturing, could benefit most from an ontology that covers this entire domain. To facilitate interoperability, the ontology could be built by aligning with a top-level and mid-level ontology and attempting to reuse the existing work as much as possible. In addition to the previous ontological work, the ontology should utilize concepts from well-established standards such as ISA-95 and ISA-88 [35,36]. A specialized ontology for continuous biopharmaceutical manufacturing could be derived from such an ontology. This ontology will also greatly help the industry, given its current desire to switch to continuous manufacturing and its lack of specialized data standards compared to batch manufacturing [37].

4.3. Applying Ontologies to Agricultural Manufacturing

According to the National Academies of Sciences, Engineering, and Medicine (NASEM) definition, bioeconomy encompasses only a part of the entire agricultural sector [25]. Specifically, it encompasses the sectors of agriculture that satisfy at least one of the following four criteria:

1. Genetic engineering is used for creating a strain or seed.
2. Advanced molecular biology techniques are utilized for marker-assisted breeding programs.
3. Large informatics databases and computational techniques are used for either breeding applications or enhanced land-use capabilities (e.g., precision agriculture, accelerated breeding, genomics-assisted planning of genetic crosses).
4. The produced biomass is used in downstream bioprocessing or fermentation process that utilizes recombinant and synthetic DNA technologies.

This ‘limitation of scope’ imposes data challenges specific to the portion of agriculture within the bioeconomy and will be the focus of the rest of this subsection. Hence detailed considerations for the entire agricultural sector are out of scope for this paper.

Genetic engineering requires understanding the impact of a genetic modification (insertion, deletion, or mutation) on the overall yield and cost of production. If the genetically modified organism is going to be used for human consumption, as indicated by Criterion 1 above, any possible health implications of the modification also need to be comprehended. Therefore, connections between production conditions and health and safety data with the genetic factors could permit a more streamlined and comprehensive analysis.

Challenges associated with Criteria 2 and 3 can be summarized as difficulties with ‘big data’ (especially genomic data), where efficient data browsing, and discovery are required. Finally, in the case of Criterion 4, the challenges are equivalent to the challenges identified in the bioindustrial manufacturing sector (Section 4.1).

Compared to the bioindustrial and biopharmaceutical sectors, the ontological landscape of agriculture as a whole seems to be more mature. Specifically, several prominent agriculture ontologies handle aspects such as agronomic techniques and variables (e.g., Agro [38]), agriculture supply chain (e.g., SCT [20]), and precision farming (e.g., ONTAgr [39]). Moreover, the agriculture sector is already using the biomedical ontologies within the OBO Foundry (e.g., gene ontology and cell ontology) to represent various cellular, biochemical, phenotypical, environmental, and molecular biology factors. Also, many application ontologies exist, such as the ontology for citrus farming [40], the ontology for crop cultivation [41], and the ontology for soil classification [42].

Keeping the outlined ‘limited scope’ in mind, the following areas of improvement might enable the current ontological ecosystem to tackle agriculture’s data challenges within the context of bioeconomy:

- To address the challenges associated with using genetic engineering, the current set of relational expressions can be further enriched and standardized to permit clear linkage of the genetic modification with process metrics and health and safety data for both animal and plant products.
- To address the challenges of using the generated biomass in biomanufacturing, a set of constructs that permit the representation of biomass in the context of biomanufacturing and the linkage of the agricultural data to biomanufacturing production data can be created.
- To ensure efficient data discovery and promote interoperability across this sector, an ontology that covers the common aspects of plants and animals can be created based on existing prominent ontologies within the respective domains.

5. LOOKING AHEAD

The analysis and rationale provided in Section 4 argue for further research and development in domain-level and application-level ontologies for biomanufacturing that are based on the top-level and mid-level ontologies discussed in Section 3. In the near-term, there are at least the following three areas related to biomanufacturing that appear promising for further exploration.

- *Ontology assisted data-driven modeling*: Recent advances in machine learning have enabled the creation of empirical models from a large amount of experimental and observational data, and the utilization of such data-driven models in several domains, including biomanufacturing [43]. Ontologies can assist in linking data-driven models (e.g., inputs, outputs, and model objective) to concepts that represent the domain (e.g., from a domain ontology for biomanufacturing, or an application ontology for upstream process control) using the hub-and-spokes approach. For example, the models deployed for process control and monitoring within the highly regulated food industry and biopharmaceutical manufacturing must be operationally transparent, and the associated data (internal model parameters and predictions) must have a high degree of traceability – these requirements can be met more effectively by the use of ontologies.
- *Ontologizing digital twin framework for manufacturing*: One of the topics of intense research in recent times is about digital twins in scientific and industrial applications. ISO has responded to part of the industrial need by releasing a series of standards on a digital twin framework for manufacturing [44]. It is very natural to examine how this framework can be used to define a domain-level ontology for digital twins in manufacturing. This can be specialized to application-level ontologies for various biomanufacturing sectors discussed in Section 4.
- *Ontological support for circular bioeconomy*: Increasing global emphasis on sustainability has given rise to the notion of a ‘circular bioeconomy’ in which the needs of circular economy are addressed using biomanufacturing [45]. Some of the sustainability benefits of biomanufacturing have been addressed already in Section 4. Domain-level and application-level ontologies that address this intersection of circular economy and bioeconomy will be beneficial to the greater global economy.

6. SUMMARY AND CONCLUDING REMARKS

The inherent complexity of biology as a science led to the application of ontology to clarify biological terms and their relationships, and to process the associated data. This proved to be very successful in increasing our current understanding of modern biology, and it also contributed to rapid developments in biotechnology. Building on this success, concerted standardization efforts undertaken over the past two decades have yielded the beginning of a hierarchical set of standards for

top-level and mid-level ontologies to cover more general cases. These ontology standards are being used now to create domain-level and application-level ontologies, especially to support the manufacturing industry.

One of the important and emerging manufacturing industry that could benefit from such ontological advances is the biomanufacturing industry (consisting of bioindustrial, biomedical, and agricultural manufacturing) that underpins the burgeoning bioeconomy. It is hoped that undertaking further research and development as outlined in this paper to advance the application of the hierarchical approach to ontologies will benefit the biomanufacturing industry immensely.

ACKNOWLEDGEMENT AND DISCLAIMER

The authors would like to thank their colleagues in standards development organizations and industrial consortia. The work presented in this paper is an official contribution of the National Institute of Standards and Technology (NIST) and not subject to copyright in the United States. Certain commercial systems are identified in this paper. Such identification does not imply recommendation or endorsement by NIST. Nor does it imply that the products identified are necessarily the best available for the purpose.

REFERENCES

- [1] ISO 23247-1. ISO 23247-1 through 4: Automation Systems and Integration - Digital Twin Framework for Manufacturing – Parts 1 through 4, International Organization for Standardization, Geneva, Switzerland (2021).
- [2] Extensible Markup Language (XML), <https://www.w3.org/XML/>. Accessed on Jan. 7, 2023.
- [3] AgGateway Open Standards Webpage, <https://www.aggateway.org/GetConnected/StandardsGuidelines/XML.aspx>. Accessed on May 15, 2023.
- [4] connectSpec Webpage, <https://oagi.org/pages/connectspec>. Accessed on May 15, 2023.
- [5] Smith, B., Ontology, in Luciano Floridi (ed.) *Blackwell Guide to the Philosophy of Computing and Information*, Oxford, pp. 155-166, 2003.
- [6] ISO/IEC 21838-1, *Information technology – Top-level ontologies (TLO), Part 1: Requirements*, International Organization for Standardization, Geneva, Switzerland, 2021.
- [7] ISO/IEC 21838-2, *Information technology – Top-level ontologies (TLO), Part 2: Basic Formal Ontology (BFO)*, International Organization for Standardization, Geneva, Switzerland, 2021.
- [8] Gene Ontology, <http://geneontology.org/docs/introduction-to-go>, Accessed Jan. 7, 2023.
- [9] Gell-Mann, M., *The Quark and the Jaguar*, W.H. Freeman and Co., New York, 1994.
- [10] Rees, M., Desert Island Discs, BBC Radio, May 4, 1997.
- [11] Nurse, P., *What is Life?*, W.H. Norton and Co., New York, 2020.

- [12] Open Biological and Biomedical Ontology Foundry, <https://obofoundry.org/>, Accessed Jan. 7, 2023.
- [13] Basic Formal Ontology, <https://basic-formal-ontology.org/>, Accessed Jan. 7, 2023.
- [14] ISO/IEC 24707, *Information technology — Common Logic (CL) — A framework for a family of logic-based languages*, Geneva, Switzerland, 2018.
- [15] ISO Standards Maintenance Portal, <https://standards.iso.org/iso-iec/21838/-2/ed-1/en/>, Accessed Jan. 7, 2023.
- [16] “An Overview of the Common Core Ontologies.”, https://www.nist.gov/system/files/documents/2021/10/14/nist-ai-rfi-cubrc_inc_004.pdf, Accessed Jan. 25, 2023.
- [17] Drobnjakovic, M., Boonserm K., Ameri, F., Will, C., Smith, B., Jones A., The Industrial Ontologies Foundry (IOF) Core Ontology, 2022.
- [18] Open Applications Group Webpage, <https://oagi.org/>, Accessed Jan. 25, 2023
- [19] Kulvatunyou, B., Ameri, F., A Supply Chain Reference Ontology Based on Basic Formal Ontology, ASME 2019 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference IDETC/CIE 2019, 2019.
- [20] Ameri, F., Wallace E., Yoder, R., Riddick, F., Enabling traceability in agri-food supply chains using an ontological approach. *Journal of Computing and Information Science in Engineering* 22, no. 5 (2022): 051002.
- [21] Hodkiewicz, M., Low, E., Woods, C., Towards a Reference Ontology for Maintenance Work Management, I-ESA Workshops, 2020.
- [22] Munira Mohd, A., Rahul, R., Otte, N., Smith, B., A product life cycle ontology for additive manufacturing, *Computers in Industry* 105 (2019): 191-203.
- [23] Kjellberg, T., von Euler-Chelpin, A., Hedlind, M., Lundgren, M., Sivard, G., Danfang, C., The machine tool model—A core part of the digital factory, *CIRP annals* 58, no. 1 (2009): 425-428.
- [24] Vlasenko, L., Lutska, N., Zaiets, N., Korobiichuk, I., Hrybkov, S., Core Ontology for Describing Production Equipment According to Intelligent Production, *Applied System Innovation* 5, no. 5 (2022): 98.
- [25] Frisvold, George B., Steven M. Moss, Andrea Hodgson, and Mary E. Maxon. "Understanding the US bioeconomy: A new definition and landscape." *Sustainability* 13, no. 4 (2021): 1627.
- [26] Menke, M., Behr, A., Rosenthal, K., Linke, D., Kockmann, N., Bornscheuer, U., Dörr, M., Development of an Ontology for Biocatalysis, *Chemie Ingenieur Technik* 94, no. 11 (2022): 1827-1835.
- [27] Grünh, J., Behr, A., Eroglu, T., Trögel, V., Rosenthal, K., Kockmann, N., From Coiled Flow Inverter to Stirred Tank Reactor—Bioprocess Development and Ontology Design, *Chemie Ingenieur Technik* (2022).
- [28] Caño de Las Heras, S., Gargalo, C., Caccavale, F., Gernaey, K., Krühne, U., NyctiDB: a non-relational bioprocesses modeling database supported by an ontology, *Frontiers in Chemical Engineering*: 118.
- [29] Process Chemistry Ontology (PROCO) github repository, <https://github.com/proco-ontology/PROCO>, Accessed Jan. 25, 2023.
- [30] Millicam, T., Jarrett, A., Young, N., Vanderwall, D., Della Corte, D., Coming of age of allotrope: proceedings from the fall 2020 allotrope connect, *Drug Discovery Today* 26, no. 8 (2021): 1922-1928.
- [31] Sarntivijai, S., Yu L., Zuoshuang X., Meehan, T., Diehl, A., Vempati, U., Schürer, S., et al., CLO: The cell line ontology, *Journal of biomedical semantics* 5, no. 1 (2014): 1-10.
- [32] Mayer, G., Montecchi-Palazzi, L., Ovelleiro, D., Jones, A., Pierre-Alain B., Deutsch, E., Chambers M., et al., The HUPO proteomics standards initiative-mass spectrometry controlled vocabulary, Database, 2013.
- [33] He, Y., Cowell, L., Diehl, A., Mobley, H.L., Bjoern P., Rутtenberg, A., Scheuermann R., et al., VO: vaccine ontology, The 1st International Conference on Biomedical Ontology (ICBO-2009): July, pp. 24-26, 2009.
- [34] Coakley, J., Hogan, N., McGuire, L., Griffin, B., Casey, C., Campbell, C., Crean, A., A methodology for knowledge management in biopharmaceutical production, *Pharm Eng* 30, no. 3 (2010): 30-40.
- [35] ISA 95 official webpage, <https://www.isa.org/standards-and-publications/isa-standards/isa-standards-committees/isa95>, Accessed: Jan. 26, 2023
- [36] ISA 88 official webpage, <https://www.isa.org/standards-and-publications/isa-standards/isa-standards-committees/isa88>, Accessed: Jan. 26, 2023
- [37] Cao, H., Srinivas M., Higgins, B., Chandrasekhar K., Fermier, A., Hausner, D., Shantenu J., Ravendra S., Marianthi I., Ramachandran, R., A systematic framework for data management and integration in a continuous pharmaceutical manufacturing processing line, *Processes* 6, no. 5 (2018): 53.
- [38] Aubert C., Buttigieg P., Laporte, M., Devare M., Arnaud E., CGIAR Agronomy Ontology, <http://purl.obolibrary.org/obo/agro.owl>, licensed under [CC BY 4.0](https://creativecommons.org/licenses/by/4.0/). 2017.
- [39] Rehman, A., Shaikh, Z., ONTAgri: scalable service oriented agriculture ontology for precision farming, International Conference on Agricultural and Biosystems Engineering (ICABE 2011), *Advances in Biomedical Engineering*, pp. 411-413. 2011.
- [40] Wang, Y., and Ying, W., Citrus ontology development based on the eight-point charter of agriculture, *Computers and electronics in agriculture* 155 (2018): 359-370.
- [41] Li, D., Li, K., Xinrong, C., Daoliang, L., Laiqing, J., Kaiyi, W., Yingyi, C., An ontology-based knowledge representation and implement method for crop cultivation standard, *Mathematical and Computer Modelling* 58, no. 3-4 (2013): 466-473.

[42] Muñoz, A., Soriano-Disla, J., Janik, L., An Ontology-Based Approach for an Efficient Selection and Classification of Soils, Intelligent Environments (Workshops), pp. 69-78. 2017.

[43] Rathore, A., Saxena, N., Thakur, G., Mishra, S., Artificial intelligence and machine learning applications in biopharmaceutical manufacturing, Trends in Biotechnology (2022).

[44] Shao, G., Frechette, S., Srinivasan, V., An analysis of the new ISO 23247 series of standards on digital twin framework for manufacturing, MSEC2023-101127, Proceedings of the ASME 2023 Manufacturing Science and Engineering Conference, June 12-16, 2023, New Brunswick, NJ, U.S.A.

[45] Carus, M., Dammer, L., The circular bioeconomy—concepts, opportunities, and limitations, *Industrial biotechnology* 14, no. 2 (2018): 83-91.