Patent Analysis Using an Ontology of Qualities of Inorganic Materials Based on Context-Dependency

Yoshinobu Kitamura[1, *,†], Koki Taniguchi[1,†], Shinichi Kato[2,†]

[1] *Ritsumeikan University, 2-150 Iwakura-cho, Ibaraki, Osaka, 567-8570, Japan*
[2] *Murata Manufacturing Co., Ltd, 617-8555, Japan*

Abstract

This paper presents the building of an ontology focused on the qualities of inorganic materials aimed at facilitating the analysis of patent documents. Acknowledging the contextual nature of qualities like "particle diameter" in inorganic materials, concepts are defined within specific contexts, such as objects. The paper discusses the application of this ontology in patent analysis, followed by the development of an ontology-based patent analysis system. This system enables the extraction of object-quality-value triples and the restoration of missing information. Through testing, the paper demonstrates the utility of this system in enhancing patent analysis processes.

Keywords

engineering, patent analysis, quality, role

1. Introduction

Efficiently analyzing patent documents is paramount for driving technological advancement. In the domain of specific electronic products within the realm of inorganic materials, on which this study primarily focuses, the number of relevant patents has exceeded 2000 annually in recent years. Despite that finding out useful information from vast quantity of patents requires efficient analysis of quality descriptions embedded in the text, existing technology does not work satisfactorily. For example, conventional statistical natural language analysis, less attention is paid to extracting numerical values written in text compared to other research fields [1], and methods using machine learning have problems with accuracy [2][3].

Moreover, a common issue arises from the omission of elements, as observed in phrases like "particles of 100 nm". In such cases, the omitted "a particle diameter" must be context-dependently inferred with "particle". Conversely, in phrases like "film of 2.0 µm", the "length" should be interpreted as "a thickness of 20 µm". These suggest the necessity of distinguishing between basic generic qualities like length and context-dependent ones like diameter or thickness as discussed in Section 2. While existing ontologies like [4][5] mainly define the

---

Proceedings of the Joint Ontology Workshops (JOWO) - Episode X: The Tukker Zomer of Ontology, and satellite events co-located with the 14th International Conference on Formal Ontology in Information Systems (FOIS 2024), July 15-19, 2024, Enschede, The Netherlands

[*] Corresponding author.

[†] These authors contributed equally.

y-kita@fc.ritsumei.ac.jp (Y. Kitamura)

CEUR Workshop Proceedings ceur-ws.org ISSN 1613-0073

former as discussed in Section 3, this paper explores an ontology specifically designed to define context-dependent qualities and the merits obtained by their utilization in patent analysis.

2. Research issues

This study addresses the following issues of qualities and terminology used to describe inorganic materials, emphasizing the need for context-dependent recognition and identification.

The first issue is to recognize relationship between synonyms; for example, between synonyms for “粒子径” “particle diameter” and “粒径” (abbreviation for “particle size.”)

The second issue involves identifying basic qualities (called **generic quality**) and context-dependent qualities, particularly in relation to objects. For instance, the generic quality “length” may be referred to as “diameter” when representing the width of a circular or spherical object. Moreover, the expression of “diameter” varies depending on the object type; for instance, “particle diameter” carries the same essence as diameter, but its expression varies based on the context in which “particle” is utilized. On the other hand, “length” may be termed “thickness” for “filmy materials.” Despite sharing a common “length value,” the terminology changes depending on the context, whether it’s referred to as diameter or thickness.

The third issue involves the utilization of the same term across different contexts. For instance, the term 粒径 (abbreviation for “particle size”), mentioned earlier, not only denotes the diameter of an individual “particle” but also refers to the average value of the particle size of an assemblage of particles in a handful of “powder”. Sometimes “particle size”（粒径）refers to the “mean of particle diameter” of “powder” or the “crystallite size” of a “crystal” resulting from the baking and hardening of particles, with potential for omitted or implied meanings. It is necessary to make these omitted meanings explicit and clearly distinguish them.

The fourth issue pertains to quality values reliant intrinsically on measurement context. For instance, “permittivity” varies based on specific parameters like frequency, temperature, and voltage during measurement. Such qualities, termed “reaction-relational qualities,” are contingent on “inputs” such as voltage applied to objects.

The research issues found in this research also include the context-dependency on manufacturing process. Although we have built an ontology of process-dependent qualities, we omit it in this paper due to space limitation.

Addressing these challenges necessitates the construction of an ontology systematically defining concepts related to properties and their representative terms (aka labels). Moreover, by explicitly delineating the contextual basis of each concept, it becomes feasible to appropriately capture the relationship between terms whose expression varies with context. Analysis suggests the need to accommodate a broad range of contexts, including object types like “particle” and “crystal,” as well as input parameters such as measurement frequency.

3. Related work and approach

3.1. Treatment of qualities and properties

The so-called “qualities” and “attributes” of entities (expressed as quality, property, attribute, etc.) are defined in various upper ontologies [6][7][8], albeit with significant differences [9]. However, delving into the intricacies of these disparities exceeds the scope of this paper; thus, readers are directed to the literature [9] for comprehensive insights. In this study, we use the YAMATO ontology [8], as also mentioned in [9], for two main reasons. Firstly, it effectively handles the distinction between **quality** and **quantity**, which is important for engineering problems. YAMATO separates quality instances from their values, allowing for clear treatment of **quality** like John's weight and its value (**quantity**. e.g., 65 kg). This differentiation ensures identical instances of this weight quality despite it denotes changes in its value over time.

The second reason is the treatment of "role concepts" and "quality roles" [8]. In YAMATO, roles are *anti-rigid*, *dynamic*, and *externally grounded* [10]. A key principle is that a **potential player** for a **role** is a **role-holder** when it actually plays the role. A **role** is an entity to be played, a **potential player** is an entity that can play a role, and a potential player becomes a **role-holder** in playing a role. In the school example, when a person (a potential role player, the class constraint of the slot expression) enrolls in a school (a context), the person plays a role of "student" in the school and becomes a student (a role-holder, role-playing entity).

In YAMATO, quality concepts are defined through a role pattern where a **generic quality** (a potential player) plays a **quality role** (a role) with respect to a specific measured **object** (a context), making it a quality role-holder [8]. A **generic quality** is the most general kind of quality and it represents basic physical parameters (e.g., length, mass, and temperature). A **quality role** is a role played by a generic quality, it includes height, weight, and body temperature. "Height" and "width" qualities of a rectangular object are both quality roles played by the "length" as the generic quality. Their representations can change based on the orientation of the object, showcasing their *dynamic* and *anti-rigid* nature, suggesting they be treated as roles. This framework, treating qualities as roles, allows for flexible definitions, particularly useful for inorganic materials. Thus, we choose to adopt YAMATO [8] and Hozo [11] for ontology development due to their shared treatment of roles.

3.2. Definitions of "generic quality" in the scientific domain

Generic quantity such as "length," "mass," and "time" in the scientific domain are rigorously defined by standards such as ISO and are specified in many ontologies such as [4][5], as well as being compared in survey papers [12][13]. In[14], an Information Model (IM) including unit conversion is established based on OM, QUDT, international standards, and other sources.

While the ontology in this study also deals with such generic qualities and units, the primary emphasis is not on their definition. Instead, the focal point centers on the context-dependency of qualities, as expounded in Section 2.

3.3. Definitions of context-dependent qualities such as diameter

Within the glossary of international standard organizations [15], generic qualities such as "length" are referred to as "quantity," while concepts like "diameter" are termed "kind of quantity" and explained as an "aspect common to mutually comparable quantities." However, the treatment of these concepts varies across various ontologies [4][16][17]. I-ADOPT[2] aims at

---

[2] InteroperAble Descriptions of Obervable Property Terminology WG (I-ADOPT WG), https://i-adopt.github.io/

interoperability among them. For instance, ontologies such as NCIT[3], PATO[4], and SIO[5] provide examples of descriptions for both generic qualities like “length” and context-dependent qualities like “diameter,” defined in parallel at the same level of the “is-a” hierarchy.

Concerning the differentiation between “diameter” and generic qualities, pioneering ontology research such as EngMath [16] has addressed this discrepancy. While OM[6] [4] defines “diameter” as a subclass of “length,” it does not explicitly articulate its relationship with objects like circles or spheres. Similarly, QUDT[7] conceptualizes “diameter” as a narrower concept of “length,” akin to “width” and “height,” without specifying its association with circles or spheres. QUDV [17] defines “width,” “diameter,” etc., as “SpecializedQuantityKind” related to the “SimpleQuantityKind” “length,” yet lacks the contexts in which they might be called differently.

In any case, qualities concepts dependent on entities like “diameter” are not clearly defined as a distinct concept type from generic qualities like “length,” while being related to both generic qualities and the types of objects (e.g., spheres or circles). In this study, we aim to define objects as contexts to address such context-dependent issue. In addition, considering the existence of non-quantitative quality such as way of birth, quality should not be modeled as kind of quantity. Following DOLCE [6] and YAMATO [8], we define quality as dependent entity rather than kind of quantity with a clear distinction between quality and quantity.

4. An ontology of qualities of inorganic materials

This section discusses a part of an ontology of qualities of inorganic materials, using concepts such as “particle” and “particle diameter,” as example and explaining related concepts.

4.1. Particle and particle diameter

The concept of “particle” is defined in Figure 1. Specialization (is-a relation) begins from the top level and progresses downward through “particular”, “independent entity”, “physical”, “continuant” (physical object), “unitary object[8]” (in accordance with YAMATO definitions), “solid object,” and “spherical object”, culminating in the subclass “particle.” The superclass “physical object” typically possesses certain qualities, thus featuring a “quality role” slot. By specifying the “diameter” slot in its subclass “spherical object” and designating its corresponding subclass “length” as a subtype of “quality,” we establish the concept that a “spherical object” possesses the quality of “length” referred to as “diameter.” In essence, within the context where “length” serves as a quality of “spherical object,” it plays the role of “diameter” (a concept of quality role. Refer to Section 3.1 and [8]), with “length” transitioning into a role-holder concept (role-playing entity) termed “diameter.” Such definitions as “quality role” based on YAMATO have advantage of uniform representation of qualities. As mentioned in Section 3.1, “height” and “width” are anti-rigid and dynamic [10] and thus quality roles.

---

[3] National Cancer Institute Thesaurus, https://bioportal.bioontology.org/ontologies/NCIT

[4] Phenotypic Quality Ontology, https://bioportal.bioontology.org/ontologies/PATO.

[5] Semanticscience Integrated Ontology, https://bioportal.bioontology.org/ontologies/SIO.

[6] The OM - Ontology of units of Measure, 2.0, https://github.com/HajoRijgersberg/OM.

[7] QUDT–quantities, units, dimensions and data types ontologies, https://www.qudt.org/

[8] Engineering processes often involve modeling objects from two inconsistent perspectives: (i) as a unitary entity, like a particle, which becomes two different particles when it is cut into half, and (ii) as amount of matter. This paper adopts the former perspective, modeling objects accordingly.

Furthermore, the specialized concept "particle" inherits the "diameter" slot, yet its role name is designated as "particle diameter." Consequently, although both terms denote the diameter of the same spherical object, the terminology (surface vocabulary) varies between particles and spherical objects. However, it is clearly defined that they essentially represent the same diameter. Additionally, using a green-colored alias node, 粒径 "particle size" is described as a synonym for "particle diameter."

Moreover, as depicted in Figure 1 under "filmy object," it is indicated that the term "length" may also be referred to as "thickness," while demonstrating that "crystal" may use the same "length" term to denote "crystallite size." This may also be referred to as "particle size."

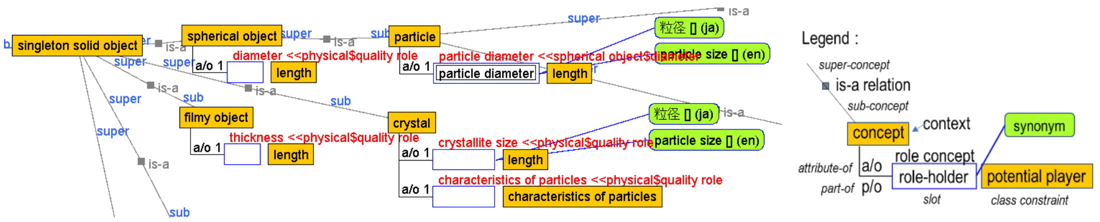

Figure 1: Definitions of "particle" and "particle diameter".

Furthermore, as depicted in Figure 2, "length" is defined as something capable of holding a quantity value. These quantity values can be expressed either quantitatively (quantitative values) or qualitatively (qualitative values). Thus, we delineate them as "length quantity" and "length qualitative value," respectively, as the class constraints of slots of "length."

In Figure 3, we defined "length quantity" as subclasses of "quality value." "Length quantity" represents the value attributed to "length," with its unit confined to "unit of length." Subclasses of "unit of length" are defined to encompass units associated with length, such as "m" and "nm." The "#" symbol preceding the class constraint "unit of length" indicates that what goes into this slot is not an instance but the class (type such as "m" and "nm") itself. If values are measured in different unit systems, two quantities, say, "10" as the number and "μm" as the unit and "0.000393" and "inch" are equivalent[9].

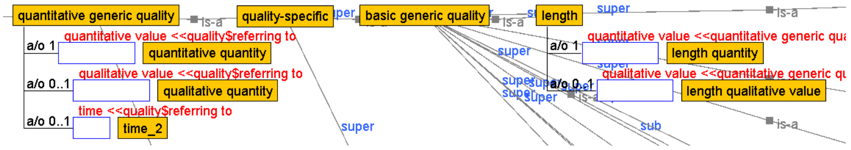

Figure 2: Definition of "length".

---

[9] Strictly speaking, these values are results of observations/measurement like in SOSA/SSN [18]. This is out of scope of this paper, though YAMATO's treatment as a representation is discussed in [19].

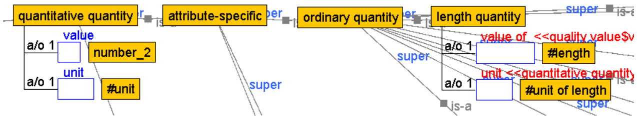

Figure 3: Definition of “length quantity”.

4.2. Reaction-relational quality

We delineate qualities such as “length” and “weight,” intrinsic to objects themselves, as distinct from qualities that manifest in response to external stimuli[10], like “permittivity” and “electrical resistance”, classified as subclasses of “reaction-relational quality”. Values of such quality “realize (manifest)” as the response to the external stimuli unlike generic qualities like “length”. To represent such reaction-relational qualities, it is imperative to reference input attribute values. We designate this as a role concept termed “excite-reaction quality”. For instance, the reaction-relational qualities “electric characteristics”, “permittivity” and “electrical resistance” depicted in Figure 4 realizes in response to specific frequencies, temperatures, and voltages. This value lacks standalone significance and necessitates contextual reference to these inputs. We define “frequency quantity” as a class constraint and designate the role concept “measurement frequency” as an excite-reaction quality of “electric characteristics.” Likewise, “measurement temperature” and “measurement voltage” are defined as excite-reaction qualities.

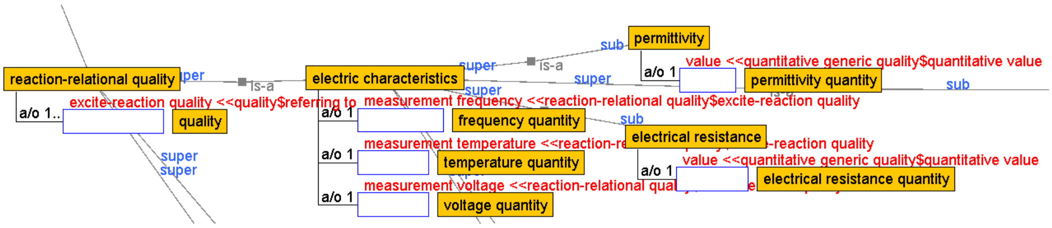

Figure 4: Definition of “reaction-relational quality”.

4.3. Qualities of a collection

The “particle” defined in Figure 1 are often treated as a collection of particles so-called “assemblage of particles”. It is often referred to as “powder”. As a collection, it is characterized by qualities such as variance, mean, median of a specific quality. As depicted in Figure 5, “variance, mean, median of particle diameter” are qualities of diameters of particles. The “mean of diameter” maybe referred to as a “particle size”.

---

[10] This is modeled as a kind of disposition in BFO [7].

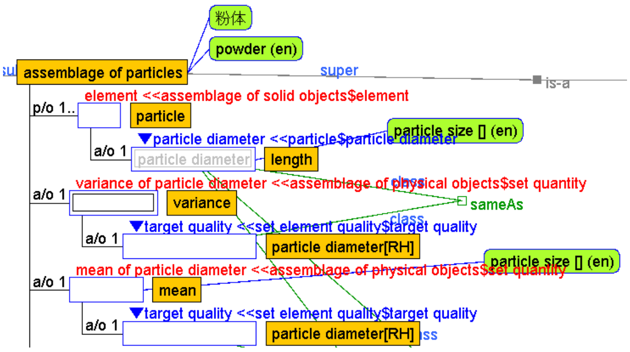

Figure 5: Definitions of “assemblage of particles (powder)” and “mean of particle diameter”.

5. Usages of the ontology for patent analysis

The ontology outlined thus far serves as the foundation for a patent document analysis system. This system is designed to extract compositions and qualities from collections of patent documents, associate them with their respective values, and store them in a database. Through the ontology, compositions and property names are standardized, accommodating paraphrases and synonyms. Moreover, the system is capable of supplementing omitted mentions of objects or quality names, establishing relationships between objects, qualities, and values whenever feasible. Further insights into the system's configuration and effectiveness are provided in Section 6.

5.1. Identification of a triplet of an object, a quantity and its value

The ontology described thus far enables us to capture the relationship between objects, qualities, and quality values (plus units) in expressions such as the following. The top line represents an example sentence found in patent documents (translated from Japanese documents), while the bottom line represents the types of concepts identifiable through analysis.

The mean of particle diameter of	cerium oxide	is 0.1-20 µm.
quality	object (powder).	quality value (+ unit)

Firstly, cerium oxide is defined as a composition under the concept of “substance”. However, it can also be interpreted as referring to the object that has it as a slot. Subsequently, as the term “mean of particle diameter” is denoted as a “mean,” it can be interpreted as a “quality of a collection,” as explicated in Section 4.3. Moreover, drawing from the definition of “assemblage of particles (powder)” presented in Figure 5, it is recognized as a “quality” signifying the mean of particle diameter within the particle assemblage. Furthermore, according to the definition outlined in Section 4.1, “particle diameter” signifies the diameter of particles and is understood to be accompanied by a length unit such as µm. Consequently, leveraging the ontology's definition, the relationship between these terms can be discerned as a triplet comprising an object, a quality, and its corresponding value. This framework underpins the discussions pertaining to issues 1-3 in Section 2.

5.2. Identification of objects from quantities

In the aforementioned example sentence, not only is the relationship recognized, but also additional information absent in the original text has been extracted. The original sentence solely denotes the composition of cerium oxide, leaving the object it represents implicit. Referring to the definition in Figure 5, which specifies that “mean of particle diameter” is a quality of an “assemblage of particles,” we can infer that it pertains to an “assemblage of particles” (“powder”) (highlighted with double underlines).

5.3. Identification of quantities from objects and quality units

In the following example sentence, the quality type “particle diameter” is augmented by considering the representation of the quality value unit (nm) and the object (particle). This augmentation is derived from the defined relationship between “particle” and “particle diameter” illustrated in Figure 1, along with the definition of “length” and “length quantity” presented in Figures 2 and 3.

particles of 100 nm → particles with a particle diameter of 100 nm
object quality quality value

On the other hand, in the subsequent example sentence, the quality type “thickness” is supplemented from the context of the “film,” as it aligns with the generic quality of length. This illustrates how distinct quality representations can be complemented based on the contextual characteristics of the objects involved.

film of 2.0 μm → filmy object with thickness of 20 μm
object quality. quality value

Consider if in the ontology both a diameter and a circumference are defined as quality roles of a particle. If the patent document has no description, it is ambiguous for quality roles. This is domain-dependent, in the inorganic material domain, by default, it refers to “diameter”.

5.4. Identification of reaction-relational qualities

The “reaction-relational quality” defined in Section 4.2 can be measured by providing specific values for “excite-reaction quality.” Hence, when presented with the following example sentence, it is inferred by convention that ε represents permittivity. Referring to the definition in Figure 4 of Section 4.2, it becomes apparent that excite-reaction qualities such as frequency, voltage, and temperature are involved in measuring permittivity. Among them, "measurement frequency" takes Hz as the unit, and 15GHz complements as the value of the measurement frequency when measuring permittivity. As a result, as shown below, 15GHz complements as the value of the measurement frequency, and 200 complements as the value of permittivity.

ε at 15 GHz is 200 ⟶ permittivity ε at measurement frequency 15 GHz is 200
reaction-relational quality value of excite-reaction quality value

6. Real usages and their effects

In this section, we discuss the implementation of systems and the effectiveness of ontologies described in the previous section.

6.1. Ontologies used in the system

We constructed an ontology to define concepts related to the qualities of inorganic materials using Hozo [11], which is exported to OWL. This ontology includes 532 comprehensive definitions of concepts such as properties and subordinate concepts within a specific domain. Additionally, elements such as titanium or barium, as well as compounds like barium titanate, are chemically classified and defined with 367 subcategories under the concept of "substance." The composition and synonyms within these subject domains are also extensively described.

6.2. Overview and procedure of text analysis

Users prepare a rough patent document set to investigate a particular area that is later processed through text analysis using our aforementioned ontology before extracting concepts and numerical ranges for storage into a relational database (RDB). Text analyses are performed by following procedure.

1. If any concept from our ontology is found within the text, it is tagged with additional augmentation like object or quality.
2. Any strings that were not tagged during step 1 are then subjected to morphological analysis using MeCab [20].
3. Units along with numbers representing values or chemical formulas are identified and assigned appropriate tags.
4. Phrases consisting of multiple concepts are tagged as a single word that includes those concepts as attributes. For example, the phrase "multilayer ceramic capacitor" contains three concepts: "multilayer structure," "ceramic," and "capacitor."
5. Connections between numerical ranges and associated concepts are identified so expressions like "using 10g particles measuring at 0.1µm" would be grouped together under one tag while completing missing information on certain abbreviated qualities if needed.
6. Information regarding objects or qualities obtained from step 5 is stored in an RDB (Relational Database), while analyzed texts themselves may also be outputted.

6.3. Examples of the complementary effect of ontology

We present an example of the effectiveness of ontology in actual searches using 139 patent documents related to Multilayer Ceramic Capacitors (MLCCs) that contain keywords such as "withstanding voltage" or "dielectric constant."

Searching for the string "粒子(particle)" resulted in 101 instances where expressions like "particles with a content greater than 90mol%" were found. However, by utilizing our ontology, we complemented the term "粒子(particle)" with its associated characteristic "粒径"(abbreviation for particle size) thereby increasing the search results to 152. As an illustrative example, consider the expression "K_2CO with a particle size below 1.0 µm." By completing this phrase with an object concept called “particle,” categorized under the quality named "particle size," we successfully included it in our search results. These completions have been verified as accurate by domain experts, and considering their completions as correct data provided by experts yield perfect accuracy and precision rate at values of both being equal to one. On the other hand, for instance, since there is no definition for “amount of warpage” within our ontology database, expressions such as “the amount of warpage in Sample One is 55 μm” failed to be complemented due to referring it back generically under “amount”. This led to failures concerning tag completion regarding “layer.” A completion recall ratio was at approximately 0.88 while also attaining an F-value close-to 0.93 (Table 1).

Next, regarding the complementation of quality and excite-reaction qualities based on their values which discussed in Section 5.4, for example, in the expression "dielectric loss at 1 MHz is less than or equal to 9.8%", we were able to identify that “dielectric loss” has a excite-reaction quality with the unit of “Hz” and could be completed as “measurement frequency”. In the aforementioned collection of patent documents, there were 191 expressions containing the term “dielectric loss”, but by searching specifically for dielectric losses at a measurement frequency of 1 MHz only, we were able to narrow down the results to 32 cases. Among these cases, there were six examples where the term “measurement frequency” was omitted. Additionally, this search also covered cases such as ranges including 1 MHz like “0.5-2 MHz”.

Furthermore, when we searched for terms such as “dielectric constant”, we obtained 265 search results. Among them, there were 26 instances where the term “measurement frequency” was explicitly mentioned and thus, it was possible to supplement it. There were two cases in which supplementing the measurement frequency failed. This happened because those patterns are currently not supported by our Japanese language processing system.

In total, we were able to accurately complete implicit expressions of measurement frequency for both dielectric loss and dielectric constant in 108 cases (True-Positive). There were 1404 cases where the frequency was not mentioned at all (True-Negative). Additionally, there were six instances where only the numerical representation of the frequency was present but could not be correctly completed as “measurement frequency” (False-Negative). Overall, summarizing these results, the accuracy was 1.00, precision was 1.00, recall rate was 0.95, and F-value was 0.97 (Table 1).

Table 1.
Summary of complementation of patent analysis.

	Accuracy	Precision	Recall	F-value
Object complementation	1.00	1.00	0.88	0.93
Quality complementation	1.00	1.00	0.95	0.97

6.4. Graphical analysis and task evaluation

After performing the tagging as described in Section 6.2, we conducted an analysis using Keygraph 0 for patent documents. First, we visualized the results with basic morphological analysis along with identification of certain chemical formulas and numerical ranges in Figure 6. As indicated by the red circles in the figure, there were instances where excessively dividing words into fragments, such as “本” (this), and chemical prefix “ジ(di)”, and divisions that lacked an object making it difficult to interpret their meaning, such as “細かい” (detailed) and “変動”

(variation). Additionally, within some islands on the graph, there was a mixture of words belonging to various contexts which made it challenging to extract any meaningful interpretation from them.

Figure 6: An example analyzed without ontology.

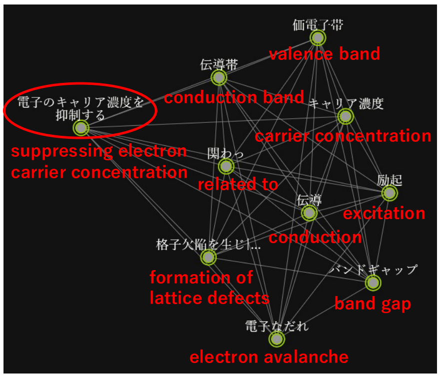

Figure 7: An example using ontology.

On the other hand, we conducted an analysis using ontology as shown in Figure 7. With the utilization of ontology, there was an increase in nodes that could be interpreted with meaningful phrases such as "suppressing carrier concentration" which made it easier to understand the meaning of islands.

By conducting such processing, it is believed that the visualization results move closer to depicting semantic relationships as a network rather than simply representing morphological relationships between tokens. As a result, domain experts can more easily evoke the effects and mechanisms discussed in patents from the visualization results.

To confirm these effects, an evaluation was conducted with three experts in inorganic materials who possessed specialized knowledge. They were given the following tasks:

1. Observe the visualization results and identify keywords related to objectives.
2. Modify and label the visualization results to improve readability of island meanings.
3. Identify key technological points described in the patents.

Table 2 shows the number of patent documents and the required time for the tasks. According to the feedback from the workers, it was reported that the analysis time was reduced to approximately 1/5 to 1/10 compared to manual analysis. Additionally, there was a tendency for less increase in required time compared to an increase in the number of patents.

Similar tasks were assigned to three different workers who provided their observations as responses. They expressed that interpreting Keygraph and adjusting calculation parameters for desired visualization results require trial-and-error processes or training efforts; however, they could understand connections between keywords and islands, and they believe it is possible to obtain necessary insights and triggers for new ideas.

After the tasks, a survey was conducted using a 5-point scale for evaluation. The average values of the evaluations from the three workers are shown in Table 3. With an average score of 3.6, it can be considered that this method is deemed useful for patent investigation. Although these results serve as reference values due to the small sample size, they also indicate variance. Evaluation items with a variance value of 0.2 had less variability in ratings.

Furthermore, regarding evaluation items with higher variances such as "Understanding overall picture of patents" and "Understanding countermeasures/solutions", interviews revealed that these were subjective and influenced by users' interpretations.

Table 2
Task time required by three workers.

Worker	Number of patents	Task time required(min) 1	2	3	sum
T	72	19	17	17	53
T	182	25	17	17	59
J	72	10	30	30	70
I	32	10	15	15	40
Mean	89.5	16.0	19.8	19.8	55.5

Table 3:
5-point evaluation of tasks by three workers.

evaluation	mean	dispersion
Analysis result	3.6	0.2
Relationship between problem and chemical composition	3.6	0.2
Overall picture of patents	4.0	0.7
Known knowledge	3.3	0.2
Unknown knowledge	3.3	0.2
Solution for problem	3.0	0.9

7. Concluding remarks

This paper has demonstrated the use of context-dependent definition of the ontology realizes complementation of missing text and thus helps the engineers capture the contents of patent documents, which has not been achieved by conventional methods. Deployment of the context-dependency on manufacturing process remains as future work. Future work also include incorporation of the LLM (Large Language Models)-based techniques.

Acknowledgements

The authors express their sincere thanks to Riichiro Mizoguchi for his valuable comments.

References

[1] Göpfert, J., Kuckertz, Weinand, J. M., Kotzur, L., Stolten, D.: Measurement Extraction with Natural Language Processing: A Review, Findings of the Association for Computational Linguistics: EMNLP 2022, pp. 2191-2215 (2022).

[2] Hao, T., Liu, H., and Weng, C.: Valx: A System for Extracting and Structuring Numeric Lab Test Comparison Statements from Text, Methods of Information in Medicine, 55(03), pp.266–275 (2016).

[3] Muffo, M., Cocco, A., Bertino, E.: Evaluating Transformer Language Models on Arithmetic Operations Using Number Decomposition, Proc. 13th Conf. on Lang. res. And Eval(LREC2022), pp. 291-297 (2022).

[4] Rijgersberg, H., van Assem, M., Top, J.: Ontology of units of measure and related concepts. Semantic Web. 4(1), 3–13 (2013).

[5] Aameri, B., Chui, C., Grüninger, M., Hahmann, T., Ru, T. : Foundational Ontologies for Units of Measure, Applied Ontology, 15(3), pp. 313-359, (2020).

[6] Borgo, S. and Masolo, C. Ontological foundations of DOLCE. In R. Poli, M. Healy and A. Kameas (Eds.), Theory and Applications of Ontology: Computer Applications, pp. 279–295. Springer. doi:10.1007/978-90-481-8847-5_13. (2010).

[7] Arp, R., Smith, B. and Spear, A.D. Building Ontologies with Basic Formal Ontology. MIT Press. (2015).

[8] Mizoguchi, R., Borgo, S., YAMATO: Yet-another more advanced top-level ontology, Applied Ontology, 17(1), pp. 211-232 (2022).

[9] Borgo, S., Galton, A., Kutz, O. (eds.), Foundational ontologies in action, Applied Ontology, 17(1), (2022).

[10] Loebe, F.: Abstract vs. social roles–Towards a general theoretical account of roles. Applied Ontology, 2 (2), 127-158 (2007).

[11] Mizoguchi, R., Sunagawa, E., Kozaki, K. and Kitamura, Y.: A model of roles within an ontology development tool: Hozo, Applied Ontology, 2(2), 159–179 (2007).

[12] Zhang X., Li, K., Zhao, C., Pan, D.: A survey on units ontologies: architecture, comparison and reuse, Program, 51(2), pp. 193-213 (2017).

[13] Keil, J.M. and Schindler, S.: Comparison and evaluation of ontologies for units of measurement. Semantic Web Journal, 10(1), 33–51. doi:10.3233/SW-180310 (2019).

[14] Martin-Recuerda, F. et al.: Revisiting Ontologies of Units of Measure for Harmonising Quantity Values – A Use Case, In Proc. of ISWC 2020, LNCS 21507, pp.551-567 (2020).

[15] Joint Committee for Guides in Metrology, International Vocabulary of Metrology, 3rd ed., https://www.bipm.org/en/committees/jc/jcgm/, (2012).

[16] Gruber, T. R., Olsen, G.R.: An ontology for engineering mathematics, Proc. of Comparison of implemented ontology, ECAI'94 Workshop, W13, pp.93-104 (1994).

[17] OMG, Systems Modeling Language, Model Library for Quantities, Units, Dimensions, and Values (QUDV), OMG Document Number: ptc/2009-08-16, pp.221-240 (2009).

[18] Janowicz, K. et al.: SOSA: A lightweight ontology for sensors, observations, samples, and actuators, Journal of Web Semantics, 56, pp.1-10, (2019).

[19] Mizoguchi, R.： YAMATO: Yet Another More Advanced Top-level Ontology, https://www.hozo.jp/onto_library/YAMATO101216.pdf (2010).

[20] Mecab: Yet Another Part-of-Speech and Morphological Analyzer, https://taku910.github.io/mecab/, (2024) (in Japanese).

[21] Llorà, X., Goldberg, D.E., Ohsawa, Y. et al.: Innovation and Creativity support via Chance Discovery, Genetic Algorithms, New Mathematics and Natural Computation, 2(1), pp.85-100 (2006).