Proceedings of the Design Society, Volume 5: ICED25
https://doi.org/10.1017/pds.2025.10305

Standards as Discourse

Jacob Collard [,1], **Eswaran Subrahmanian** [,2,✉], **Spencer Breiner** [,1] and **Ram D. Sriram** [,1]

[1] *National Institute of Standards and Technology,* [2] *Carnegie Mellon University*

✉ es3e@cs.cmu.edu

ABSTRACT Technical standards provide order and consistency in application domains; however, standards development organizations produce large families of related documents containing significant amounts of information that can be difficult to access, evaluate, and produce consistently. We describe standards as linguistically, socially, and conceptually dynamic constructs using theory drawn from systems engineering and linguistics to create a model of standards documents that can be updated, evaluated, and queried to retrieve information reliably. We describe the theoretical basis for this model from multiple perspectives and explain broadly how it can be used to retrieve relevant information from standards.

KEYWORDS: standards development, natural language processing, speech act theory, formal models

1. Introduction

Technical standards are documents that provide order and consistency in particular application domains and are critical in industrial and economic development, national security and safety, and competitive advantage. Standards in the information age are also crucial in coordinating network effects across vertically integrated industrial complexes. Despite the importance of standards and their evolution in technological change and innovation, standards are commonly understood as static entities that remain relatively stable with some periodic revisions over time once agreed upon by standards developers. However, especially in the case of information technology standards, changes to standards may occur constantly and rapidly (Egyedi and Blind, 2008). Our primary hypothesis is based on the observations and work on the role of discursive practices in the collaboration in sociology and management where in a group working towards a common goal and, in this case, standards, the conversational practices identify the language and resources in establishing a collective identity (Hardy et al., 2005). In the case of standards, the common ground created by the standard is also tied to the product and identity of the group developing them. Hence, by examining and codifying the discursive practices of conversations embodied in text, memos, emails, speech acts documents, and other material artifacts, we can track and trace the evolution of the linguistic common ground in the making (Hardy et al., 2005). There are two main themes to this paper. Our first objective is to characterize the dynamics of standards development processes. We describe the standards process as a linguistically and conceptually dynamic socio-technical process with a set of discourse actions used to achieve precision in their specification. There are multiple dynamics of change for writing and rewriting standards, taking place at different levels of implementation and use.
Our second objective is to illustrate the use of linguistic discourse structures in standards to create a framework for developing computational support for the representation and development of standards. Discourse structures are representations of linguistic interactions above the sentence level. There are two linguistic discourse arenas in standards development: the discourse of developing the standard and the discourse between the standards developers and consumers via the standard itself. We do not address the first arena, though we expect there to be some similarities between the two. The second type of discourse includes particular stylistic patterns for creating definitions, requirements, and references to other standards. The International Standards Organization (ISO) and other standards development

organizations (SDOs) have templates and style guides intended to control the language used in standards to be more precise (International Organization for Standardization, 2021). Some standards are also available in precise, machine-readable languages such as XML (eXtensible Markup Language) and UML (unified modeling language), with the Standards Tag Suite XML schema being published in 2017. However, most standards still follow a relatively simple document model.
In effect, understanding standards development as a dynamic process necessitates new approaches that will scale in enhancing the participation of many stakeholders from academia, industry, and government through traditional voluntary and industry-led processes. In effect, standards are developed by a social network of technical professionals, industry and government representatives, and any other relevant stakeholders.
The rest of the paper is organized as follows: Section 2 reviews the basics of standards and the characterization of dynamics of standards development. Section 3 describes the theoretical and problematic basis for supporting the dynamics of collective standards development and Section 6 will describe the future work.

2. Background

In a general sense, a standard may be viewed as "a document established by consensus and approved by a recognized body, that provides, for common and repeated use, rules, guidelines, or characteristics for activities of their results, achieving optimal degree of order in a given context" (International Organization for Standardization, 2024). The reference to an approved body covers cases conducted under the auspices of recognized national and international standards bodies.
The main reason that standards, especially IT standards, have proliferated over the last 100 years is the economic benefit of standards accrued through reduced transaction costs and enhanced network effects (Kindleberger, 1982). Reduction in transaction costs leads to economic efficiency, increased market size, reduced barriers to entry in the market, and expanded competition. While stable standards reduce transaction costs, changes in standards are imperative due to changes in technology, innovation, and intellectual property issues (Egyedi and Blind, 2008). There is a duality between order and change that is not conflicting but complementary in contrast to organizational management literature that has traditionally viewed them as exclusive (Brunsson et al., 2012). With this perspective of the dynamics of standards as mechanisms that bring order while accommodating change to account for innovation and learning from standards use, this paper focuses on defining the models and mechanisms that could help create support structures for standards.
Standards have the image of stability over time. This imagery comes from standards created for physical artifacts and their components in the last century, but it is not valid in the general case. Standard interface specifications enable increasing modularity, interoperability, and interchangeability requirements to address mass production while managing variety across different markets. Reducing uncertainty in quality and uniformity in production practices also requires standards for consistency in requirements specification, production, and assembly. Despite the need for order that standards aspire to, they evolve as more precise measurement instruments and needs are discovered. Standards, technological change, and innovation are closely intertwined in terms of participation by firms, users, and other interested parties (Foray, 1994). The benefits are derived by firms in terms of investor confidence in the product and technology, expansion of the market, and potential differentiation, leading to incremental innovation. In contrast, at times, suppressing radical innovation (Foray, 1994). The effects of standards on technological change to create order in the market, compatibility with other standards, and interoperability directly impact the potential for network effects in an industry and economic development. Technologically advanced countries are anchored in their ability to create standards to ensure economic growth through complementary and innovative products, resulting in incremental and radical innovation capabilities. The importance of managing standards and technological change is at the heart of economic development, as well as the market power exercised globally by a firm or a country (Shapiro and Varian, 1999).
Standards development is a slow process of consensus and cooperation between stakeholders, as in SDOs. The SDOs have developed procedures and methods for developing standards with various criteria and negotiation processes, both specification and creating consensus. Figure 1 below shows a simple sketch of the process of standardization, starting from specification, maintenance, feedback from implementation, and the succession of standards (Egyedi and Blind, 2008)
This simple model describes the continuous dynamic nature of standards evolution, which is the process of specific types of socio-linguistic discourse by participants that leads to agreements and revisions over

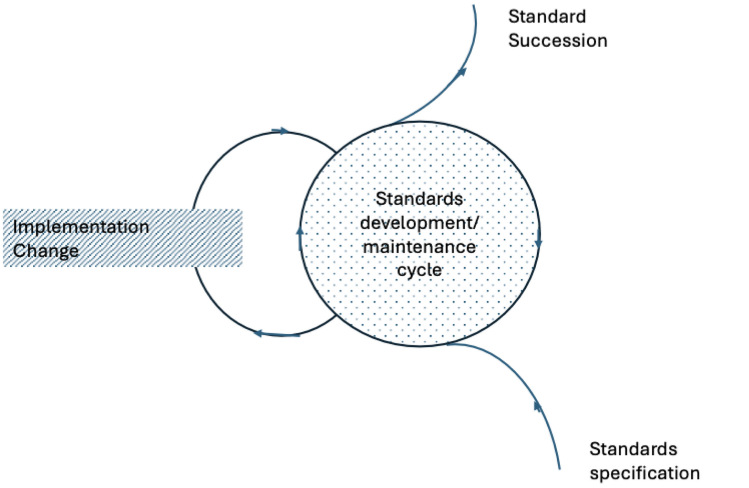

Figure 1. Standards development and revision life cycle (adapted from Egyedi and Blind (2008))

long periods of use of a specific standard in a domain application. Figure 1, above, points to the evolution of standards from drafts to a process of identification of standards to revision to convergence in their use and diffusion for International Accounting Standards (IAS) over 30 years. The details of the process of evolution from national standards to global use of IAS are described in detail by Botzem and Dobusch (2012).

Egyedi and Sherif (2008) identify potential sources of changes to the context of the definition and use the standards:

- "A flaw in the concept or content of the standard, e.g., because the scope is too wide and standards become too comprehensive and unworkable;
- An aspect of the standards process, e.g., lack of consensus; absence of an important stakeholder;
- The quality of the standard specification, e.g., ambiguous terminology, errors or omissions; and
- The way the standard is implemented, e.g. partial implementation due to cost-constraints."

There are different classes of change that may occur throughout the standard's life cycle: some implementation changes will have significant impact on the standards ecosystem, while others (such as fixing a typo) will have minimal impact. Standards processes also have document status, configuration management, and voting procedures for acceptance of changes, which implement the model in Figure 2.

3. Supporting dynamic standards development with NLP and knowledge structures

The need for computational support for standard-setting is critical due to the proliferation of new standards, mainly because of the increased complexity of information structures of standards that are partially organized to coordinate the stakeholder language but insufficient to address the problem of maintaining semantic information structures.

However, the modeling of the standards still needs to be formal and often poses issues for the changing landscape of standards. The goal of this paper is to address the usually neglected process of support for standards development despite existing tools that minimally support the versioning of standards as documents rather than changes that are more granular, reflecting changes in terms and procedures and their interrelationships to other terms and procedures in different sections of standards. This is a common state of affairs across all types of standards despite their origin and level of participation by different stakeholders. Another goal of this paper is to create a formal discourse model of standards that can be marked up conceptually to create computable standards that can serve as a basis for future smart standards development processes using natural language processing and other AI techniques to create terminological maps of relationships between fundamental concept definitions, actual implementation details, and translation for diffusion in practice. AI and systems modeling for standards are needed to synthesize useful information from the vast number of dynamic standards in the industry today. However, AI systems for standards must be interpretable and tunable to ensure that the information retrieved is accurate and maintains the precise nature of technical standards. If implemented correctly, AI will vastly improve access to information in standards and reduce participation costs in standards development and application.

3.1. Dynamic Semantics

Dynamic semantics is a family of pragmatic theories that attempt to account for inter- and intra-sentential context changes when determining the meaning of an utterance. Many linguistic theories make use of dynamic semantics, including discourse representation theory (DRT) (Kamp, 1981), file change semantics (Heim, 1982), and dynamic predicate logic (Groenendijk and Stokhof, 1991). Many forms of dynamic semantics implement an update semantics (Veltman, 1996), in which each statement made by an interlocutor causes an update to a representation of the state or context stored by the interlocutors. In an update semantics, there is usually a global representation of *discourse context* or *common ground*, often called C. Each utterance in a discourse *updates* the discourse context according to a particular function, determined in part by the utterance's semantics and the content. One way of formulating this intuition is with the formula in Equation 1. In this model, the discourse context at time t is represented by C_t, which is calculated from the context at t - 1 via an update function U and the utterance u_t made at time t.

$$C_t = U(C_{t-1}, u_t), t \geq 1 \quad (1)$$

Though this generic formula applies to many theories of dynamic semantics, it contains several elements which vary from theory to theory. In particular, the discourse context C can be represented in many different ways, though it is commonly construed as the set of possible worlds - those worlds which are consistent with what the interlocutors have agreed to accept as true. The update function U can also be defined and constrained in many ways, as can the representation of an individual utterance u.

If the discourse context is a set of possible worlds, then simple declarative sentences representing assertions can also be represented as sets of possible worlds, and $U(c, u)$ can be defined as $C \cap [[u]]$, where $[[u]]$ represents the semantic content of u - in this case, the set of worlds in which u is true.

The key insight of update semantics means that each utterance contributes to the global context in some way, and that semantics can make reference to this information as needed. This allows dynamic semantics to account for complex pragmatic phenomena such as anaphora (coreferential terms depending on context) and presupposition (inferred content which must be true in order for an utterance to be understood, regardless of its value). Presupposition in particular can be handled by placing certain constraints on the context before an update can be successfully completed (Kasper et al., 1999).

3.2. Speech Act Theory

Speech Act Theory (Austin, 1962; Searle, 1969) is a theory describing the pragmatic (context-dependent) nature of utterances in natural language. In Speech Act Theory, an utterance typically belongs to one of several different speech acts types. This allows different utterances to play different roles in conversation: some are intended to convey information, while others may be intended to bring about a change in the interlocutor or commit the speaker to a certain course of action.

Searle's initial taxonomy of speech acts included the following types (Searle, 1969):

- Assertives, in which the speaker conveys information that they commit to the truth of;
- Directives, in which the speaker directs an interlocutor to undertake a course of action;
- Commissives, in which the speaker commits themself to a particular course of action;
- Expressives, in which the speaker describes their emotional state;
- Declarations, which perform a particular action by being performed.

Speech acts are often described in terms of how they relate *word* - the conventional content of the utterance - to the *world* - the reality in which the interlocutors find themselves. In the above taxonomy, for example, assertives and expressives generally use words to describe the world, while directives, commissives, and declaratives use words to change the state of the world. This is referred to as *direction of fit* - word to world or world to word.

Speech act theory of this type has been applied to industry domains. For example, it has been used to model office interactions (Auramaki and Lyytinen, 1988) and human-computer interactions (Winograd, 1986).

Roberts (2018) provides a more recent refinement and simplification of Searle's taxonomy. In her model, there are only three speech act types: **assertions, questions**, and **directions.** There is a correspondence between these three speech act types and the three common sentential moods (indicative, interrogative, and imperative), but it is not a strict correspondence. While interrogative sentences typically correspond to questions, sometimes they can correspond to other speech act types, such as directions. For example, though phrased as an interrogative, the sentence "Could you pass the salt?" generally constitutes a direction rather than a question.

This work also provides a more formal model for the interpretation of speech acts, which combines speech act theory with dynamic semantics. Each of her three speech act types corresponds to a different component of a tripartite representation of the discourse context.

An assertion, when accepted, updates the *common ground* (CG), which is a set of possible worlds. Each assertion further restricts this set based on its semantic content.

A question, when accepted, updates the *questions under discussion* (QUD). The QUD are a stack of questions such that the answer to each question on the stack provides a partial answer to the question below it on the stack. Questions can be represented in different ways, but are commonly represented as a partition over the set of possible worlds, where each equivalence class is a possible answer to the question.

A direction, when accepted, updates the *domain goals*, which are contextualized goals for each interlocutor. A direction only affects the domain goals for a single interlocutor, and each domain goal is associated with a particular set of circumstances under which it should be completed. The discourse context in this model also contains additional information, including the full set of interlocutors, the full set of utterances made in the discourse so far, and a total ordering of the set of utterances. This can be formalized such that C is a tuple $\langle I, G, M, <, CG, QUD \rangle$ where:

- I is a set of interlocutors
- $G_i \in G$ where $I \in i$ is the set of i's publicly evident desiderata, which are themselves maps from sets of worlds to sets of worlds, the former representing the circumstances and the latter representing the desired world state;
- M is the set of all utterances made thus far in the discourse;
- $<$ is a total ordering over M storing the chronological order of utterances;
- CG is the common ground, a set of possible worlds
- QUD is the questions under discussion, a stack of questions.

Other possible formalizations amount to more or less the same basic idea, but the somewhat underspecified version above should suffice for the purposes of this paper.

When an utterance is made and processed by the interlocutors, they decide whether to accept it together. The mechanics of this are not made explicit but can be observed in conversation: expressions such as "okay" or "sure" and feedback like nodding can indicate the (temporary) acceptance of an utterance, while expressions such as "wait" or "what?" can indicate a lack of acceptance. Once an utterance has been accepted, the discourse is updated depending on the speech act type of the utterance. If a question is accepted, it is added to the top of QUD. If an assertion is made, the asserted proposition is intersected with CG. If a direction is made, it is added to the listeners set of circumstantial desiderata.

There are a few key takeaways from this model:

- Speech acts and dynamic semantics can interact in a formal, structural model.
- The update function of a dynamic semantics can be determined, in part, by the speech of the utterance.
- Semantic content and speech act are related, but not directly - the same speech act type may have different types of semantic content, depending on the context.A discourse can be modeled as a structured object and updated based on utterances in context.

3.3. Standards

There are many different types of standards, which have different structures.

The ISO and the International Electrotechnical Commission (IEC) have together formalized some of the typographical and stylistic features that standards documents should have; these are published in the ISO/IEC Directives, Part 2 (International Organization for Standardization, 2021). This document describes

linguistic and structural features that should be included in ISO/IEC standards documents, though many of these directives are also borrowed by other standards bodies.
Standards are typically described in terms of their elements. Elements are linguistic or structural objects within a standard, and may include sentences, clauses, a term and its definition, or a figure, though the term is never formally defined. Elements are, however, clearly distinguished in terms of their roles: there can be *normative* elements and *informative* elements. Normative elements describe the scope of the document and set out provisions, which are expressions that convey information, criteria for conformance, and possible courses of action. Informative elements assist in understanding and use of the document and its context.
The ISO/IEC Directives prescribe specific language for certain element types. For example, *requirements* are expressed using the modal *shall* - "The packaging shall not transmit any odour or flavour to the product and shall not contain substances which can damage the product or constitute a health risk." In order to claim compliance to a standard, it must be the case that the packaging does not transmit any odor or flavour. Though this is intuitively true, it is important to distinguish this usage of *shall* from other, similar modals such as *should*, which is used in standards to indicate a recommendation and *not* a requirement.
Standards are also structured in particular ways. A document is primarily divided into clauses (sections), subclauses, and paragraphs. Some elements are mandatory while others are optional. Terms and definitions have special formatting to indicate the relation between the *preferred term*, any *admitted terms* (accepted synonyms), and *deprecated terms*.
In 2017, the National Information Standards Organization (NISO) published a standard defining the Standards Tag Suite (National Information Standards Organization, 2017). This standard provides an XML schema for the publication and exchange of standards, including full-text and metadata. STS contains tags which make much of the structure of a standard explicit and machine-readable. Structural elements such as the abstract, annexes, tables, definitions, and references are all represented in STS. However, some semantic information cannot be expressed directly with STS. For example, the differences between provision types (requirements, recommendations, statements, and instructions) are not made explicit in STS XML. The more general distinction between normative and non-normative content is sometimes made explicit, but only for certain elements. Furthermore, not all documents written in STS make use of the full tag set; STS authors can choose how granular they would like their document description to be.
Though there has been significant adoption of NISO STS, most standards are still only available as PDF files. Even when standards have been developed using XML, the PDF distribution is sometimes considered more accessible, though it is less useful for interoperability and analysis. Using STS can make the automated recognition of specific element types easier and more accurate.

4. Towards the Formalization of Standards

The previous sections suggest two main points about standards:

1. Standards are dynamic.
2. Standards are structured.

By the first point, we mean that standards change over time. During the development of a standard, this change is extremely rapid. Over the lifecycle of a standard, the rate of change may slow, but it never stops. Some standards are updated regularly according to a planned schedule, while others are only updated when necessity or new developments dictate. Even physical standards need to be updated as new technologies and research methods come into play.
By the second point, we mean that standards contain both explicit and implicit structures. The hierarchical relationships between documents, clauses, and sentences are all explicit to human readers, though they may not be apparent to a machine depending on the standard's format. The relation between a term and its definition may be difficult to recover from a PDF. However, standards also contain implicit structures, which are not even represented in course-grained STS. The semantic relationship between a term and each instance of its use is an implicit structure.
If standards are dynamic and structured, a formalism for representing them should support their dynamism and structure. We propose a high-level framework for standards that meets these basic requirements. There are many possible formalisms that could specify the representation of a collection of

standards; what we describe is just one example that instantiates the formal representation of standards in the context of these desiderata.

4.1. Goals

Our goal in this paper is to provide a framework for formalizing standards that better support their dynamic structures. This formalization has three broad, practical applications:

- The formalization allows for the automatic detection of errors.
- The formalization allows new presentations of information.
- The formalization allows for fine-grained version control and dependency management.

These applications are dependent on having a structured and dynamic view of standards. Without explicit structure, error detection may still be possible (e.g., with AI), but interpretable descriptions of those errors become easier to guarantee using explicit, formal structures. Dynamic error detection means that errors can be detected during standards development, thus making the standards development process more efficient. Flat documents such as PDFs can be presented in many ways, but converting between presentations is not always straightforward. The presentation of information goes beyond formatting and includes exactly what information is presented to which users in what order. By presenting information dynamically, new changes do not require the acquisition of an entirely new document; instead, the new relevant information can be queried independently. Version control is, by its nature, tied to dynamics. Still, without structuring the information, version control can only apply at the document level. Without the representation of informational dependencies, tracking what information is impacted by a change becomes difficult.

4.2. Standards as Discourse

We propose that the representation of standards can follow many of the same principles as the representation of discourse in linguistics. This may initially seem unintuitive; standards are documents and not conversations, they do not have clear interlocutors, and they are relatively static. However, considering the goals of representing standards in a structured, dynamic way, discourse models are reasonably well-suited for standards. Standards are primarily linguistic objects. Though they do contain non-linguistic information, most standards are primarily text written in natural language. When standards change dynamically, this can be thought of as a discourse update. However, this is not the only discourse update in standards. In fact, many of the structures in standards can be represented as discourse updates that occur in a particular order. In natural language, discourse is always presented chronologically. In standards, there may be other orders, but there *is* still an order to the information. Discourse models are ultimately representations of semantic structures. This is needed to make standards structured and dynamic in a useful way.

A formal model of standards as discourse needs the following components : a model of discourse context C, a model of an utterance S, an update function $U : C \times S \rightarrow C$ and a method for generating utterances. In the following sections, we will discuss how each of these components may be implemented for standards. We do not claim that this describes the only possible implementation, but suggest that by implementing standards in this way, it becomes possible to achieve the goals outlined in the previous section. There are certainly other ways to achieve these goals, but by using linguistic models of discourse theory we can leverage pre-existing work in the linguistics literature.

4.3. Standards Discourse Context

For the purpose of representing standards, we propose a model of discourse context which simplifies and generalizes previous work on discourse. Because standards contain several types of explicit, interrelated structures, we suggest that a graph be used to represent the discourse context. Updates to the discourse context can add, remove, or modify nodes and edges.

This results in a simple, more general interpretation of discourse context than that provided in Roberts 2018. Instead of having three separate components for the context set, questions under discussion, and domain goals, different roles are represented by different nodes in the graph and the relations between them. The actual information presented in a discourse model of standards will differ somewhat as well, and it is possible to represent many different discourse models in graph form.

The specific representations of different kinds of utterance in standards and how they are stored in the graph will be discussed in Sections 4.4 and 4.5.

4.4. Standards Utterances

In spoken language, an utterance is a continuous unit of discourse. In theories such as Roberts 2018, an utterance is the unit that carries illocutionary force (i.e., a speech act). Something similar should be the case for standards. Intuitively, an utterance in a standard is usually a sentence. However, there are certain edge cases that should be considered.

4.4.1. Clauses

Clauses in standards serve the same role as sections and subsections in journal articles and other scientific documents: they provide the document with hierarchical structure and inform the reader about the purpose or topic of a particular portion of the document. Clauses are of particular interest from a discourse perspective because they are hierarchical; models that only consider sentences as a unit of discourse may not be able to handle this hierarchical structure. However, it is possible to flatten this structure by considering the clause heading to be its own, simple discourse unit which introduces the clause's title and other metadata, and then introducing relations between the clause and all of the other utterances within it.

4.4.2. Terminology

Terminology in standards is typically presented in a very structured way, with a clear delineation between the term to be defined, the definition, and other metadata related to the term's usage in the standard. In a discourse, we consider each element of a definition to be its own utterance (though not all are propositions); we then define discourse relations between the definiendum and the other components of the definition, which are subordinate to the definiendum. In this way, the relationship between all of the elements is maintained, but each element also stands alone, in case it needs to be targeted as a single utterance. For example, it may be useful to isolate a proposition in a note attached to a definition.

4.4.3. References

References in standards typically refer to other standards documents. They may appear inline or in a special references section. References are often somewhat underspecified. Though they always describe the standard that is being referenced, they do not always describe specific elements in that standard that are used in the current document. Unfortunately, additional information may not be available in current standards documents, but it may be possible to infer usage information through some of the techniques described in this paper. For now, however, a reference as discourse simply contains the identifier corresponding to the referenced document.

4.5. Standards Updates

The update function for standards viewed as discourse is relatively simple. Recall that the discourse model consists of a graph. Nodes can represent elements in a standard, and edges represent relations between them. Each update simply adds the appropriate nodes and edges corresponding to the utterance type. For sentences, the update function adds the sentence as a node, with the text of the sentence as its label. A relation can then be added from the sentence node to the clause containing it. Clauses introduce a new node, with the clause title as their label. A relation is also added between the clause and its immediate parent, if it has one. Terminology introduces a new node for each element of the definition, as described above. It also introduces a new edge between the definiendum and each other element. References introduce a new node with the label corresponding to the identifier of the referenced document. If the discourse is meant to represent multiple standards documents, then the reference node can be linked directly to the referenced element.

5. Operations on Standards Discourse

5.1. Version Control

Version control is a system for tracking changes made to data. ISO standards are currently versioned according to the year of publication. The version applies to the entire standard, and when standards make reference to other standards, they typically reference a specific version. This creates dependency relations between standards based on the version of each standard: if a change is made to one standard, other standards which reference it must either be updated or continue to reference the outdated standard. However, it may be the case that a standard only needs to reference one part of another. When the referenced standard is updated, the update may not involve the actual referenced content. Some updates to standards may require changes to downstream standards, while others will not. In such cases, it would ideally be possible to track versions not by documents, but by more fine-grained elements of a standard, or even by individual propositions. In discourse, interlocutors may choose to accept or decline specific propositions into the common ground; the same should be true in standards. Standards developers should be able to make changes to individual propositions and identify the downstream changes in the common ground (the model of the standard) automatically.

5.2. Information Extraction

When standards are presented as flat documents, it is difficult to extract specific types of information. If documents are encoded only as PDFs, then extracting information is even more difficult, as structural elements including sections and paragraphs may not be contiguous.

There are many different types of information that a user may wish to extract from a standard. Representing a standard as a discourse model will provide new methods of information extraction for different use cases. There are many possibilities here, as this is a very broad category. Some examples include: extracting normative content from a standard, extracting content containing a particular term, or extracting content which makes reference to other standards (or extracting the set of related standards).

5.3. Error Detection

Certain linguistic, stylistic, and logical errors may occur in standards that can be identified automatically. The simplest such error to identify is a circular definition, in which one concept is ultimately defined in terms of itself. Such cycles may be of varying lengths, spanning two or more definitions, and may be the result of stylistic or logical problems with one or more of the definitions. By detecting them automatically, standards organizations can quickly identify potential problems with a standard and find effective solutions.

6. Future Work

This paper has described a framework for representing standards using discourse models. It does not provide an actual implementation of this model, nor the results of applying that model to real standards. This implementation is left to further work, as there are many possible ways to implement this framework. Natural language processing (NLP) offers many such techniques to arrive at the formal model by analyzing syntactic, semantic, and pragmatic properties of text.

This framework provides tools not only for improving the standards development and dissemination of standards, but also for the study of standards content. By extracting information from standards, researchers can determine stylistic patterns which can in turn be used to develop new operations on this model of standards and find new applications.

Finally, this model involves the representation of standards. But the development of standards introduces its own discourse that is in many ways separate from standards themselves. Encoding the discourse of standards development could further improve the process.

7. Disclaimer

Certain commercial entities, equipment, or materials may be identified in this document in order to describe an experimental procedure or concept adequately. Such identification is not intended to imply recommendation or endorsement by the National Institute of Standards and Technology, nor is it intended to imply that the entities, materials, or equipment are necessarily the best available for the purpose.

References

Auramaki, E. and Lyytinen, K. (1988). A speech-act-based office modeling approach. *ACM Transactions on Information Systems*, 6 (2):126–152.

Austin, J. L. (1962). *How to do things with words*. Harvard University Press, Cambridge, MA.

Botzem, S. and Dobusch, L. (2012). Standardization cycles: A process perspective on the formation and diffusion of transnational standards. *Organization Studies*, 33 (5-6):737–762.

Brunsson, N., Rasche, A., and Seidl, D. (2012). The dynamics of standardization: Three perspectives on standards in organization studies. *Organization studies*, 33 (5-6):613–632.

Egyedi, T. M. and Blind, K. (2008). *The dynamics ofstandards*. Edward Elgar Publishing.

Egyedi, T. M. and Sherif, M. H. (2008). Standards' dynamics through an innovation lens: Next generation ethernet networks. In *2008 First ITU-T Kaleidoscope Academic Conference-Innovations in NGN: Future Network and Services*, pages 127–134. IEEE.

Foray, D. (1994). Users, standards and the economics of coalitions and committees. *Information economics and policy*, 6 (3-4): 269–293.

Groenendijk, J. and Stokhof, M. (1991). Dynamic predicate logic. *Linguistics and Philosophy*, (14):39–100.

Hardy, C., Lawrence, T. B., and Grant, D. (2005). Discourse and collaboration: The role of conversations and collective identity. *Academy of management review*, 30 (1):58–77.

Heim, I. (1982). *The Semantics of Definite and Indefinite Noun Phrases*. PhD thesis, University of Massachusetts, Amherst.

International Organization for Standardization (2021). ISO/IEC directives, part 2 - principles and rules for the structure and drafting of ISO and IEC documents. Technical report, International Organization for Standardization, Switzerland.

International Organization for Standardization (2024). Consumers and standards: Partnership for a better world. Accessed = 2024-12-5.

Kamp, H. (1981). *A theory of truth and semantic representation, volume 135 of Mathematical Centre Tracts*, pages 277–322. Mathematisch Centrum, Amsterdam.

Kasper, R. T., Davis, P. C., and Roberts, C. (1999). An integrated approach to reference and presupposition resolution. In *Proceedings of the ACL '99 Workshop on the Relationship between Discourse/Dialogue Structure and Reference*, College Park, Maryland. Association for Computational Linguistics.

Kindleberger, C. P. (1982). *Standards as public, collective and private goods*. IIES.

National Information Standards Organization (2017). STS: Standards tag suite. Technical report, National Information Standards Organization, Baltimore, Maryland, USA.

Roberts, C. (2018). *Speech acts in discourse context*. Oxford University Press.

Searle, J. (1969). *Speech Acts*. Cambridge University Press.

Shapiro, C. and Varian, H. R. (1999). *Information rules: A strategic guide to the network economy*. Harvard Business Press.

Veltman, F. (1996). Defaults in update semantics. *Journal of Philosophical Logic*, 25 (3).

Winograd, T. (1986). *Understanding computers and congition: a new foundation for design*. Norwood, NJ.