INTERNATIONAL COMPUTATIONAL LAW JOURNAL

CODEX SOURCE ARCHIVE

Drafting X2RL: A Semantic Regulatory Machine-Readable Format

Patrick A. McLaughlin, Walter Stover

Computational Law Report · Release 3.0 · 2021-05-14

Reader edition prepared by The Research University under CC BY 4.0. This cover was added; the original PDF pages follow unchanged. Original authorship, publication identity and source links are preserved. This independent republication is not sponsored or endorsed by Stanford University, Stanford Law School, CodeX or MIT.

Volume 1 · Issue 5 · Evidence, Interpretation and Machine Judgment

Original source and license evidence are recorded on the article landing page.

MIT Computational Law Report

Drafting X2RL: A Semantic Regulatory Machine-Readable Format

Patrick A. McLaughlin, Walter Stover

Published on: May 14, 2021

License: Creative Commons Attribution 4.0 International License (CC-BY 4.0)

ABSTRACT

> *We introduce a new regulatory machine-readable format called X2RL. Existing "syntactic" formats for presenting regulatory text primarily focus on reducing management costs of legal bodies and regulatory agencies. In contrast, the economic costs of legislation manifest primarily as the cost of understanding and complying with the body of regulations and associated informal documents that implement and enforce this legislation. A machine-readable format that focuses not only on the internal structure of documents, but also on their content, meaning, and external structure, will help reduce both the management and economic costs of legislation and regulation. X2RL is a new "semantic" machine-readable format that reduces these costs by enhancing regulatory documents and other formal and informal legal documents with rich metadata fields that inform machine and human readers of the types of effects a document will have, relevant industries and agencies that the document applies to, and how restrictive a particular document is.*

---

1. Introduction

Attempting to understand the vast body of regulation in the United States is a rather byzantine task. Using a Shannon entropy benchmark, many financial regulations and parts of the U.S. Code rank higher in linguistic complexity than Shakespeare's plays (McLaughlin et al. 2020). In 2011, it took 70,000 pages of instructions to explain the federal tax code (Davis 2017). The corpus of federal regulations alone would have required over three years for someone to read through in their entirety as of 2016 and has been consistently growing since then (McLaughlin et al. 2017). This leads to two primary costs: the cost of management of the production of law, and the cost of compliance by private entities.

Research suggests that, when considering opportunity cost, the economic costs of compliance may be significant. Coffey et al. (2020) estimate that the growth in regulation has dampened economic growth by an average of 0.8 percent per year since 1980, for a total price tag of $4 trillion by 2012. This dampening occurs primarily through restrictions and requirements that prohibit certain operations or mandate additional costs to businesses in various forms such as compliance and reporting requirements, distorting and sometimes deterring the business investments that drive long-run economic growth. Theoretically, and in practice, any services, tools, or

technologies that assist the layperson in understanding and complying with the regulatory body reduce these costs—hence the thriving industries of RegTech and FinTech, and popular software such as TurboTax.

We do not yet have a TurboTax for all regulation (or even just all federal regulation, setting aside state and local regulation), in large part because of the sheer size and complexity of the formal and informal regulatory body. One of the frictions that contributes to this difficulty of comprehending the regulatory body consists in the natural language form of the documents. While these documents are now electronically available for download in a variety of formats, they lack built-in machine-readable metadata that assist with parsing the original natural language of the document. In lieu of this kind of metadata, machines must still rely almost entirely on natural language processing (NLP) techniques to extract any meaningful information.

NLP approaches have already demonstrated value added to understanding the regulatory corpus and lowering its economic cost. The Mercatus Center's QuantGov project, for example, collects all federal and state regulatory documents in the United States and several other countries, and applies simple NLP algorithms to parse them for a set of words and phrases that are likely to restrict an individual's or business's legal choice, such as "shall" and "must." Another set of machine learning algorithms maps regulations to the industries that the regulations are most likely to affect. By applying these algorithms to the entire body of regulatory text, users can gain a macro perspective on the number of restrictions present in the regulatory corpus, as well as what particular states, industries, and agencies have the most restrictions.

However, these NLP algorithms have their limitations, because they are still programmatically parsing only natural language. In the process of digitization as described in Wong (2020), pulling the text from the overlying HTML layer or using Optical Character Recognition (OCR) technology to extract it from a scanned image remains at the lowest level of digitization. An alternative is to move away from a purely natural language format towards writing legislation and regulations in a *machine-readable* format. Machine-readable formats, such as XML or JSON, represent a higher level of digitization, and enable greater understanding of documents by machines through the use of ontologies, elements, attributes, and tags that create metadata that describe the structure and content of text in a useful way to machine readers. This enrichment of metadata constitutes a partial move towards a "Rules as Code" paradigm where documents, though not fully machine-consumable in the sense of being able to be executed directly by a machine in some manner, can allow for greater

comprehension and deeper analysis by machines of the body of legislative documents ("Cracking the Code: Rulemaking for Humans and Machines"). Machines will no longer process only natural language but will also parse metadata that can communicate meaningful information about the structure and content of that natural language. This, in turn, can enable development of tools and programs that will help alleviate both legislative management costs as well as the economic costs of compliance.

Two of the most prominent forms of machine-readable formats for regulatory documents consist of the USLM XML format for U.S. regulations, and the Akoma Ntoso standard (also XML) for international regulatory documents. The U.S. currently offers downloads of documents in the Electronic Code of Federal Regulations in XML markdown format with a hierarchy of tags and attributes. Similar hierarchies are present in the Akoma Ntoso standard, which has been used to good effect in cases such as translating UN documents automatically between languages, or programmatically generating visualizations of data contained in UN General Resolutions.

However, both of these formats primarily benefit the drafters and other legal experts directly involved with producing, organizing, and amending the body of legislative and regulatory documents. As machine-readable formats, both USLM and Akoma Ntoso emphasize improving legibility of the *internal structure* of legal documents, rather than their *content*. Moreover, both USLM and Akoma Ntoso lack sufficiently developed methods of connecting the body of legal documents together in a machine-readable way. In other words, the current paradigm of legislative machine-readable formats focuses on reducing legislative management costs, not on the economic costs of legislation, specifically those who must comply with the body of regulations.

In the following section, we provide a brief overview of USLM and Akoma Ntoso, before describing a new differentiation between "syntactic" machine-readable formats that focus only on legislative management costs and "semantic" formats that could also help reduce the economic costs of complying with regulatory and other legal restrictions. We then move into a technical proposal for a new semantic machine-readable format, X2RL, and explain how this new semantic format will extend beyond previous languages by reducing these economic costs through new, richer metadata primarily concerning the content and meaning of formal and informal regulatory documents that enforce and implement the primary legislation. We will also discuss how X2RL knits this wider universe of documents together by enhancing these

documents' external structure through stronger referencing mechanisms and other metadata. We follow this by discussing some use cases for X2RL and conclude by discussing the current state of X2RL's development, obstacles, and next steps.

2. Existing Machine-Readable Regulation Formats

Within the United States, there are two dominant types of machine-readable formats for regulation: USLM and Akoma Ntoso. USLM, or United States Legislative Markup, is an XML standard prepared by the Office of Law Revision Counsel (OLRC) that consists of an abstract model generalizable to bills, resolutions, statutes, and other legislative materials, as well as a concrete model specifically fitted to the U.S. Code. The abstract model consists of three sets of elements: the *primitive set* that serve as the fundamental building blocks of the model, the *core set* that defines a basic document model and contains the most granular set of elements used for structuring a document, and the *generic set* that mostly imports basic elements from XHTML such as header, row, column, and paragraph tags.

Additionally, USLM has a set of *attributes* that offer additional information on the elements that make up the abstract model. These attributes serve the primary functions of *identification*, *classification*, *annotation*, *description*, *action*, *amending*, *linking*, *value-holding*, *dating*, *versioning*, *cells*, and *notes*.

According to the USLM User's Guide, the USLM format has three primary goals:

1. Ease of Learning
2. Extensibility
3. Editability

All three goals relate to the OLRC's decision to design USLM for maximum flexibility, acknowledging the significant changes the U.S. legislative format has undergone in the past two centuries, the subsequent changes that will follow in the future, as well as the existence of "anomalous structures," such as federal statutes that "… do not conform to any commonly accepted drafting style, past or present" ("USLM User's Guide", p.15). In addition to downloading a selected subset of the U.S. code in XML format, the OLRC also maintains an HTML version of the code translated from the XML.

In addition to USLM, the UN and Organization for the Advancement of Structured Information Standards (OASIS) maintain an international XML schema known as Akoma Ntoso (meaning "linked hearts"). Akoma Ntoso is broader and more comprehensive than USLM, likely because it must attempt to encompass the legislative

and regulatory processes of all the world's countries, and not just the United States. Rather than one basic document model as in USLM, Akoma Ntoso has six basic document structures, and twelve actual document types (due to some document types sharing their document structure with other document types). Like USLM, Akoma Ntoso categorizes a document's hierarchy into upper and lower levels.

3. Semantic vs. Syntactic Machine-Readable Formats

While differing in many ways, USLM and Akoma Ntoso share a focus on detailing the internal structure of legislation. The core set of the USLM consists of elements that demarcate preambles, headings, tables of contents, the type of legal document, dates, and appendices. Additional elements specifically help those who must edit the document, including the **<instruction>** and **<action>** elements that describe amendments to legislation and what changes specifically need to be made. Universal IDs and dates, meanwhile, help keep track of what version of a document they are working with. These types of elements are very useful for drafters, editors, and anyone else directly involved with producing, amending, or keeping track of the body of legislation. Indeed, this appears to be the primary audience for both USLM and Akoma Ntoso. Monica Palmirani, a prominent member of the Akoma Ntoso project, specifically highlights how machine-readable formats could enhance three processes used to reduce the costs of legislative management: consolidation, codification, and recasting (Palmirani 2011).

As the USLM and Akoma Ntoso focuses on the internal structure of a document, we consider these examples of *syntactic* machine-readable formats to allude to the standard meaning of syntactic as focusing on structure, rather than content. While valuable to the legislative drafting and amendment processes, a syntactic machine-readable format possesses certain limitations in other respects. First, by focusing on internal structure, a syntactic format may lack elements that enable robust, informative linking between documents, limiting cross-communication across the body of documents, and thus also limiting the value and learning that could be gained from machines who can walk across this structure. As a result, a syntactic format will result in a less cohesive and structured body of documents, especially when considering both primary legislation as well as the subsidiary documents and regulations that derive from the primary legislation—not to mention the extensive body of informal guidance and recommendations that regulators often issue that end up acting as best practices for end users. Akoma Ntoso does include elements to handle both internal (intra-document) and external (inter-document) references, with the goal of allowing readers

to navigate between documents, but to our knowledge, machine parsers cannot reliably use these external references to read documents outside of the immediate document in a comprehensive and structured way that is informative to its users.

Second, a syntactic markup format lacks elements that focus on the *content* of a document's text, which limits the richness of datasets that machines can generate and the accessibility of legislative and regulatory documents to the general public. For instance, while USLM has a **<level>** tag with attributes identifying what level a particular unit of text occupies within the document, it does not have attributes describing the function of that unit of text, such as whether it is directly restrictive on any action, or if it grants powers, or presents a formula for calculating some relevant function. This is representative of both USLM and Akoma Ntoso containing elements that prioritize reducing costs of legislative management. However, they both lack content-rich metadata crucial to also reducing the economic costs of legislation and regulation.

While a syntactic format undoubtedly aids the legislative process, we believe that a machine-readable format that focuses not just on the internal structure, but also on the *content* and *external structure* of regulatory, legislative, and informal documents could provide additional value to lawmakers, researchers, and the private citizens who derive privileges from and must comply with the existing body of regulation. We call this content-focused, external-oriented alternative a *semantic* machine-readable format. By connecting regulation with the primary legislation and other associated and supporting documents through robust linking elements, a semantic format enables machines to programmatically walk through all levels of the body of legislative and regulatory text in a structured manner that more clearly relates documents to each other. This could improve communication between machines and these documents, and consequently make it easier for people—including the legislators and regulators themselves—to build better tools for understanding these texts.

Additionally, by focusing on content, a semantic regulatory markup language provides richer data on legal restrictions found in regulatory and other documents, making it easier to categorize, aggregate, and parse. As an example, a semantic format could group regulations by industry, by agency, or by the topic they concern. This greatly enhances the utility of existing NLP techniques, which can now more accurately categorize parts of text into relevant sub-groups to analyze. These more nuanced tools will help improve the accessibility of legislation and regulation, assisting with comprehension and compliance with legal restrictions. Moreover, by tagging parts of

documents that directly restrict private action, a semantic format can also allow us to identify which regulations or laws are most restrictive and enable more precise regulatory reforms. In short, this could reduce the economic friction of legislation and regulation, and also enable new tools that benefit legislators and regulators, such as the ability to dynamically change attributes of parts of regulations that may be turned “on” or “off” during an emergency response.

Of course, there is no such thing as a pure semantic or syntactic format – USLM and Akoma Ntoso both allow for and contain some amount of semantic metadata. Our proposed format will retain many of the syntactic metadata elements of these languages. By calling USLM and Akoma Ntoso syntactic formats, we simply mean that the *focus* is on describing the *structure* of a legal document, rather than its *content*. Such focus on syntactic structure surely assists administrators and drafters of legal documents, but it does not describe the meaning of documents or link them together in ways that might be more useful for the end users of these documents. However, a semantic machine-readable format that includes these features may now be feasible, given recently developed computational techniques for parsing large amounts of text. Such a semantic approach would make it easier for people to execute their plans than it was with a purely natural language format, as well as assist regulators in analyzing and managing the existing body of regulation. Below, we detail a semantic machine-readable format for legislation and regulations that includes these features.

4. X2RL

With the semantic vs. syntactic format distinction in mind, we will now detail a new machine-readable format for regulatory and other legal documents. This format is currently under research in the Policy Analytics Team at the Mercatus Center, with a tentative name of X2RL (eXtensible Regulatory Reporting Language)**.** The X2RL format is designed to be fully compatible with Akoma Ntoso; it will follow the same hierarchical structure as detailed in the Akoma Ntoso User’s Guide, but with the addition of new attributes, tags, and models designed to enrich the metadata of the document’s content. None of these additions directly contradict Akoma Ntoso’s requirements, and so X2RL and Akoma Ntoso should be able to dock relatively easily.

Here are the primary goals of X2RL:

1. Rich metadata concerning the content, intent, scope, and meaning of legislative and regulatory documents.
2. Depth of external structure linking together documents.

3. Efficient data interchange formatting to reduce cost of machine reading and consumption.

The following specifications should be understood as directed at the *abstract* rather than the concrete level. In other words, we want to present here a high-level overview of the X2RL format, including primarily the addition of new basic tags and elements.

To accomplish the first goal, X2RL builds primarily from Akoma Ntoso by adding elements that enhance metadata concerning the *content* of the text, rather than just its structure; by adding elements that enhance the overall *external* structure of the regulatory body, documents are more concretely connected. X2RL creates content-rich, semantic metadata through three primary features:

1. A new container for text, **<provision>**, which acts as the basic container for any document fragment that directly mandates or restricts certain actions.
2. New attributes for the provision container with semantic metadata about the *effect* of a given unit of regulatory text (e.g., granting of a power, restricting actions) and the *direction* of a given effect (i.e., restricting actions of private entities, or internal, restrictions on an agency or other government entity).
3. Expansion of the ontology to include new elements describing compliance and enforcement dimensions to legislative documents, as well as conditional clauses and exceptions or exemptions to provisions.

X2RL accomplishes the second goal of strengthening external structure through the addition of elements and attributes that clearly identify the “parent” regulations or legislation that give legal authority to the current document, or which the current document cites, and the “children” regulations or legislation that the current document gives legal authority to, or which cite the current document. This essentially creates a living “bibliography” for each legislative or regulatory document that enhances both the external horizontal and vertical structure in the legal corpus through the referencing of existing documents to one another. This bibliography allows for machines to programmatically determine the structure of the legal corpus, enabling them to “walk” between documents in any direction, including vertically. In turn, this enables tools that can tell you exactly which document is the most authoritative, for example, as the machine works through the relevant metadata. These elements will also expand on Akoma Ntoso’s existing reference mechanisms through the addition of references to external sources that are not directly cited within the document but have relevance to the document’s content. For instance, a regulation

may contain references to guidance and recommendations issued by the relevant agency that serve as “rules of the road” for interpreting that regulation.

For an example of an X2RL-formatted regulation, Appendix 1 contains two examples of X2RL provision text containers with attribute metadata identifying the effect and direction of the regulatory text, as well as associated NAICS industry codes and whether the provision is currently active. This sample also includes a tagging element, **<compliance>**, from the X2RL ontology, which wraps around an “Eligibility Certificate” mentioned in this document. The tagging element contains a @refersTo attribute linking the tag back to its parent Top-Level Class, **<TLCCompliance>**, which contains attribute metadata on the monetary and time costs of compliance (obtaining the certificate) and any associated targets. In this case, the target is “eligibleVessel,” another entity in the ontology for this document.

To accomplish the third goal, we have selected JSON as the canonical form for X2RL rather than XML. This is due to its stronger compatibility with interoperable features, such as REST APIs. A JSON standard will not come at the cost of graphical representation, as both XML and JSON can be translated into HTML. However, adopting a JSON standard may hinder interoperability between X2RL and both Akoma Ntoso and USLM, which use XML. While this may increase the costs of bridging between X2RL and these languages, the advantage of JSON is in the reduction of processing costs. The workload of displaying and retrieving documents can be done on the client side rather than the server side. A JSON-based format in the long run will make it easier, less expensive, and faster to deliver machine-readable documents, particularly in combination with machine-executable functions.

Nevertheless, a JSON format comes with its own limitations. First, JSON lacks a built-in notion of attributes, making it more difficult to distinguish proper content from metadata. JSON also lacks the infrastructural maturity to validate the integrity of documents with XML schemas and document type definitions, as well as convert them from the canonical form into other formats. This contrasts with XML’s capacity to convert between XSLT and XSL-FO. XML has this infrastructure because it was designed as a long-term storage format, whereas JSON was primarily designed as a data interchange format. However, we believe that JSON’s design for more efficient data interchange outweighs these costs when considering a machine-readable format for widespread and frequent machine reading and consumption, especially coupled with X2RL’s rich external structure permitting wide aggregation of documents.

We anticipate that this expanded abstract model detailed above can be customized to some degree to fit specific bodies of regulation, including different agencies or states. However, we do want to make X2RL a more restrictive format than Akoma Ntoso or USLM, with additions to the hierarchical structure and ontology needing approval from an X2RL-affiliated committee. This incurs the penalty of less interoperability between jurisdictions and we will need to take into consideration how robust a more restrictive format will be to legislative changes in the future. We will certainly need to strike a careful balance. However, we favor increased restrictiveness because too flexible a format will degrade the quality of machine aggregation and analysis, which will make it harder to achieve our goals of using machine readers to help comprehend the legislative and regulatory body. Custom names for the same ontological entity could impair analyses aggregating compliance and enforcement actions relevant to that entity.

5. Use Cases

In this section, we will briefly detail hypothetical use cases. These cases are extrapolated in part from the QuantGov team's experience in applying natural language algorithms to existing corpuses of natural language regulatory documents.

5.1 Finding relevant documents

X2RL-formatted documents provide richer fields of metadata for indexing in a searching solution (e.g. an ElasticSearch domain). A search domain using X2RL elements will allow for different kinds of searches currently not possible or only possible with intensive manual effort. For example, in the process of searching by industry NAICS code, by topic, or filtering by the number of restrictions in each regulation.

5.2 Reading an individual document

As with USLM or Akoma Ntoso, X2RL-formatted documents will have a corresponding HTML display for easy reading on a digital interface or for a print version. Unlike USLM or Akoma Ntoso, however, the finer-grained structure of X2RL will enhance the accessibility of the document. Some examples include the possibility of highlighting which clauses are direct restrictions on external entities as opposed to internal, or identifying which ones are restrictive on a particular industry related to the observer. This can be shown through a popup of a NAICS code attribute when a relevant section

heading is clicked. Readers could even isolate specific sections of the text by filtering based on X2RL tags or attributes.

5.3 Aggregating relevant document for analysis

An individual or organization with relevant software expertise could aggregate X2RL-formatted documents, perhaps through web scraping or through direct API access to the relevant U.S. government office. Consider an entrepreneur interested in starting a small business in the medical technology sector. They query a government API for all state regulations with NAICS code identifiers for the relevant sector, in order to identify which states have the least stringent restrictions. Depending on the level of machine-executability available in the X2RL format, they may even be able to arrive at a rough estimate of the annual compliance costs and time requirements by state without requiring additional legal assistance.

5.4 Clarifying localized definitions and requirements

Constructing a semantic web of legal documents can help with clarifying terms and requirements that vary in their instantiation across jurisdictions. The Environmental Protection Agency (EPA), as an example, may issue regulations requiring licensing to handle asbestos in a household environment. However, the EPA may defer to state authority to determine what these licensing requirements look like. A person or company searching these licensing requirements would ordinarily have to manually check for licensing requirements in their state; however, with X2RL-formatted documents, the regulation, itself, could contain the necessary metadata to point the user to the proper state requirements. Alternatively, a term used in the regulations of one state may differ in interpretation in another state, a difference that X2RL-formatted documents can help account for programmatically.

6. Conclusion

While many technical details remain under active consideration, our intent here is to spark a conversation on what kind of machine-readable format is optimal for legislation and regulation, and how choosing a semantic over syntactic format may provide more benefit in the long run. While the existing formats of USLM and Akoma Ntoso have undoubtedly already improved the accessibility of documents, they have primarily focused on improving the machine-readability of the *internal structure* of the document. In contrast, we believe that a semantic format such as X2RL can also make the content and intent of legislation and regulation more legible to machines. This will simplify finding, accessing, analyzing, and aggregating regulatory documents. We

propose that the X2RL format can help alleviate not only legislative and regulatory management costs, but also the overall economic cost to society by reducing the friction introduced through the accumulation of specialized natural language documents of legislation and regulation.

Of course, many decisions will have to be made before X2RL goes from abstract idea to concrete solution. For one, we need to confirm the interoperability issues of choosing JSON as the canonical format for X2RL documents as opposed to XML, as we do want to maximize the compatibility between X2RL, Akoma Ntoso, and other existing formats. With regards to the actual additions of elements that operationalize X2RL, we believe that these are technically simple enough that they can be added onto the Akoma Ntoso format without requiring significant reworking. We are not changing how either format describes their internal structure, but are merely proposing adding additional information into these existing schemas.

Another issue is the *role* of any machine-readable format, especially an enhanced one such as X2RL that actively seeks to improve the legibility and accessibility of regulations to laypersons. As Ma (2020) asks, is a machine-readable translation simply a coded version of the regulation, or does it have its own legal authority? When interpreting a document, we must expect the attributes and elements of X2RL to become active in these considerations. Would the attribute **<type: restriction>** take precedence over any other alternative interpretation that might argue that that specific clause is *not* a restriction? Though we may freely translate from the X2RL-formatted document to a natural language guise in HTML, we cannot ignore the underlying elements structuring the document. This will invariably involve a discussion on whether the natural language of the text itself is the primary form of the document, or whether the machine-readable format is now the legal standard.

While we do believe that eventually machine-readable documents could be beneficial as the primary form of legislation, and that X2RL could serve as the foundation for that eventual stage, that does not need to be the case for this format to improve on existing ones. Although the emphasis on content rather than structure will inevitably involve more interpretation, we see X2RL as serving to enact existing legislation and communicate it more clearly through a precisely-defined technology that more clearly represents its restrictions in the form of regulations and other documents, akin to the way traffic lights concisely communicate the content of existing traffic laws (Ma et al. 2020). While the machine-readable copy of a regulation may not be strictly binding in the legal sense, it can attempt to more precisely demarcate or circumscribe the intent

and limits of a regulation at the time of creation, assisting later interpretation of the scope of a legislative or regulatory document.

Bibliography

Coffey, Bentley, Patrick A. McLaughlin, and Pietro Peretto. “The Cumulative Cost of Regulations.” *Review of Economic Dynamics* 38 (October 1, 2020): 1–21. https://doi.org/10.1016/j.red.2020.03.004.

Davis, Steven J. “Regulatory Complexity and Policy Uncertainty: Headwinds of Our Own Making.” SSRN Scholarly Paper. Rochester, NY: Social Science Research Network, April 29, 2017. https://doi.org/10.2139/ssrn.2723980.

Ma, Megan. “Writing in Sign: Code as the Next Contract Language? ✍ □ □.” *MIT Computational Law Report*, August 14, 2020. https://law.mit.edu/pub/writinginsign/release/1.

Ma, Megan, Dmitriy Podkopaev, Avalon Campbell-Cousins, and Adam Nicholas. “Deconstructing Legal Text: Object-Oriented Design in Legal Adjudication.” *MIT Computational Law Report*, November 20, 2020. https://law.mit.edu/pub/deconstructinglegaltext/release/1.

McLaughlin, Patrick A., Jonathan Nelson, Jim Pagels, and Oliver Sherouse. “The Impossibility of Comprehending, or Even Reading, All Federal Regulations.” Mercatus Center, October 19, 2017. https://www.mercatus.org/publications/regulation/impossibility-comprehending-or-even-reading-all-federal-regulations.

McLaughlin, Patrick A., Oliver Sherouse, Mark Febrizio, and M. Scott King. “Is Dodd-Frank the Biggest Law Ever?” *Journal of Financial Regulation.* Advance online publication, February 5, 2021.. https://doi.org/10.1093/jfr/fjab001.

Vergottini, Grant. “The Pocket Guide to Akoma Ntoso.” Xcential Corporation, April 2018. https://akomantoso.io/resources/

Wong, Meng Weng. “Rules as Code: Seven Levels of Digitisation,” March 27, 2020.

“Cracking the Code: Rulemaking for Humans and Machines.” Accessed January 31, 2021. https://www.oecd.org/innovation/cracking-the-code-3afe6ba5-en.htm.

"United States Legislative Markup User Guide." Office of the Law Revision Counsel, U.S. House of Representatives, July 2013. http://xml.house.gov/schemas/uslm/1.0/USLM-User-Guide.pdf.

Appendix 1.

```
<TLCCompliance id="eligibilityCertification" href="/ontology/enforcement/x2rl/certification/eligibilityCertification"
cost_monetary="$600", cost_time="1mo" targets="eligibleVessel" showAs="Eligibility Certification"/>


<provision effect="restriction" direction="external" NAICS="{NAICS:[21, 23, 4711]}" active="true">
        <p>  Requirement.—A vessel may not operate under section 3 unless
                the vessel has received an
                <compliance refersTo="#eligibilityCertification"> Eligibility Certification.</compliance>
        </p>
</provision>

<provision effect="power" direction="internal" active="true" required="true">
        <p>  Issuance.—The Administrator of the Maritime Administration of
                the Department of Transportation is authorized to issue an
                <compliance refersTo="#eligibilityCertification"> Eligibility Certification
                </compliance> with respect to any vessel that the Administrator
                determines, in his or her sole discretion, meets the requirements set
                forth in section 2(4).
        </p>
</provision>
```

X2RL-Formatted Regulation