Xalgorithms Foundation

Response to FINRA's Request for Public Comment on the Future of Its Machine-Readable Rulebook Initiative

Dr. Joseph Potvin, DBA (Project Management), MPhil (Applied Economics)
Executive Director, Xalgorithms Foundation
50 Hines Road, Suite 240, Ottawa, Ontario, Canada K2K 2M5

February 21, 2023

The Xalgorithms community greatly appreciates this opportunity to comment on FINRA's Machine-Readable Rulebook Initiative.

To: Jennifer Piorko Mitchell, Office of the Deputy Corporate Secretary,
and, Alex Khachaturian, Director, Office of Financial Innovation
Financial Industry Regulatory Authority (FINRA)
1735 K Street, NW, Washington, DC 20006-1506

1. About Us: Xalgorithms Foundation and the "Data With Direction Specification" (DWDS)

The "Data With Direction Specification" (DWDS) enables a distributed, general purpose method for any person to author, publish, discover, fetch, scrutinize, prioritize and optionally automate normative assertions (rules) on informatic networks between rule-makers and rule-takers (individuals, organizations and/or machines), via any application and platform, with precision, simplicity, scale, volume and speed. The intended general-purpose capability is referred to as "an Internet of Rules".

Xalgorithms Foundation Inc. is a Canadian not-for-profit (No. 953777-5) operating globally. Xalgorithms Working Groups[1] can involve businesses, governments, academics and not-for-profits, civil society communities and individuals. Each Working Group is subject to its own written charter, and any funds donated to it are managed by the Working Group, not the Foundation. To avoid competition with commercial firms that it seeks as contributors, Xalgorithms' bylaws prevent it from engaging in fee-based contracting or licensing, and instead it is resourced solely through financial donations and in-kind contributions.

Xalgorithms Foundation's sole purpose is to convene organizational and individual contributors to the collaborative design, implementation and use of the "Data With Direction Specification" (DWDS), of which the present author, Joseph Potvin, is the lead designer.[2] The co-founder and primary angel donor to Xalgorithms, William Olders, is President and Chief Technology Officer of DataKinetics a private firm that for several decades has provided a high volume transaction processing solution called "tableBASE" to several of the world's largest banks, credit card companies, insurance firms, equities traders, and other types of financial services.[3] Whereas tableBASE runs on centralized mainframe platforms under proprietary licensing, its tabular declarative in-memory method for transaction rules

1 Xalgorithms Working Groups are roughly modelled on the structure and practises of the Internet Engineering Task Force (IETF) and its Working Groups: https://ietf.org/how/wgs.

2 Potvin, J. (2023). *Data with Direction: Design Research Leading to a System Specification for 'an Internet of Rules'. Submitted to the Board of Examiners in partial fulfillment of a Doctorate in Business Administration—DBA (Project Management).* Université du Québec—Outaouais Campus (UQO)]. *License CC-by 4.0*
https://gitlab.com/xalgorithms-alliance/data-with-direction-specification/dwds-documents/-/tree/master/current

discovery and processing was redesigned from the ground-up by Potvin as a distributed decentralized solution, with guidance from Wayne Cunneyworth, retired Chief Data Scientist at DataKinetics. DWDS was implemented through the coordination and contributions of Don Kelly, manager of the 'developer environments' team at Shopify. The resulting end-to-end system is simpler, more efficient and more flexible than the original centralized design, while retaining an emphasis on speed and volume. Core contributors to the current DWDS reference implementations include Huda Hussain, a student of Interactive Multimedia & Design, and Ted Kim, a full-time data scientist with the Government of Canada. Calvin Hutcheon, a creative technologist recently graduated from the Maryland Institute of Art (MICA), has contributed to the user interface design elegance and interaction. The Xalgorithms community operates under a 100% free/libre/open source model based on the Apache 2.0 license (RuleMaker Web app and RuleTaker embedded component) and the AGPL 3.0 license (RuleReserve network service), while documentation and data are shared under the CC-by 4.0 International license.

The DWDS RuleMaker, RuleReserve and RuleTaker are together designed to enable context-sensitive event-triggered transmission of normative data (expressions that contain MUST, MAY and SHOULD, and their synonyms and negatives) when they are:

- 'in effect' for given dates/times, identities and jurisdictions, are
- 'applicable' to a set of industry and product/service categories, and are
- 'invoked' by particular event circumstances.

2. Our Perception of the Context and Purpose of FINRA's Machine-Readable Rulebook

The 850 "FINRA Rules" and thousands of interpretative texts, policy statements, change notices, and other guidance documents produced by the Financial Industry Regulatory Authority (FINRA), shape the self-regulated market operations of more than 3,500 brokerage firms and over 600,000 registered securities representatives in the United States.[4] FINRA's regulatory role is exercised under delegated authority from the Securities and Exchange Commission (SEC), in addition to the SEC's own legal framework of statutes (Securities; Securities Exchange; Trust Indenture; Sarbanes-Oxley; Dodd-Frank; JOBS) and operational rules, regulations, schedules and interpretations.[5] The member-funded not-for-profit FINRA, and the government-funded SEC statutory agency, jointly pursue securities market integrity and investor protection.

With operating revenues in excess of USD$1B annually, FINRA's activities include oversight services (consolidated rule-making, surveillance, examinations, fraud detection, enforcement, dispute resolution); membership (applications, registrations, training, certification, communication); and transparency services (automated reporting, advanced data analytics, verification audits, and formal examinations of securities firms based on risk, scale and scope of operations).[6] FINRA's in-house capability for informatics and data science involves over 500 software developers, who currently have more than 100 software applications under management. They enable the organization to maintain continuous surveillance of securities market activity, processing approximately 6 terabytes of data per day running hundreds of surveillance algorithms on an average of a billion financial transaction events to detect patterns that may signal market manipulation, insider trading and many other unfair activities

3 TableBASE is a mainframe-based system capable of checking millions of transactions per second against an in-memory data store with 4,500 tables of digital rules. Globally in 2023 about 6 billion transactions per day are validated on the various implementations of tableBASE.

4 FINRA was established in 2007 through consolidation of the member regulatory functions of the 'National Association of Securities Dealers LLC' and 'NYSE Regulation LLC', a subsidiary of the New York Stock Exchange. Through contracts it also took on responsibility for regulating the 'Nasdaq Stock Market', the 'American Stock Exchange', and the 'International Securities Exchange'. It would consolidate regulatory rules and enforcement, and operate utilities for trade reporting and essential over-the-counter operations.

5 https://www.sec.gov/about/laws/secrulesregs

6 In 2021 FINRA referred 758 fraud and insider trading cases for prosecution, suspended or barred 655 individual traders, expelled or suspended 4 brokerage firms, imposed $130M in fines, and ordered $47 in victim restitution.

in stock and bond markets. The empirical results of this scrutiny inform disciplinary actions such as censures, fines, suspensions, expulsions, and restitution to harmed investors.

In mid-2018 FINRA launched consultation and developmental work on a taxonomy for an eventual machine-readable rulebook. A two-level categorization of regulatory and industry terms was produced, with summary themes and a hierarchy of detailed topics. In Autumn 2022 FINRA then launched two online prototype services with an initial set of the 40 most frequently viewed rules from full collection of 850. A Web application called FIRST (FINRA Rulebook Search Tool) provides an interface for users to locate FINRA rules through a step-wise selection of categories. And the FINRA API Platform (Application Programming Interface) facilitates automated keyword queries of the database of this sample of rules.

3. Three Questions Relating to Future Development of FINRA's Machine-Readable Rulebook

The present submission is in response to FINRA's "Request for Comment on Its Machine-Readable Rulebook Initiative" announced on 21 October, 2022. That request lists numerous questions, but only three of those under the section title 'Future Development' will be addressed here, following which we table and respond to two questions of our own.

[D.2] Are there readily available applications and methods that FINRA should consider to automate the identification and application of its taxonomy to its rulebook?"

In early February 2023 the following mini-experiment was conducted with a generic 'Generative Pre-trained Transformer' (GPT):[7]

```
1. Write your own short XML semantic schema for the domain of 'financial securities regulation'. Include 'Condition' and 'Action' elements to express the logic structure within any particular regulation.
```

The application produced a roughly acceptable XML Schema Definition (XSD), equivalent to what a junior-to-intermediate developer might prepare.

The application was then tested on how well it could automate the application of its own generated XSD to the unaltered text of a sample FINRA rule.

```
2. Apply your FinancialSecuritiesRegulation semantic schema to the following regulation, with standard XML syntax, and using an n-tuple to express the logic structure (not a sequence of nested elements) in order to accommodate multiple condition permutations: "FINRA Rule 4140. Audit. (a) FINRA may at any time, ..."
```

This prompt generated a reasonably good XML representation of FINRA Rule 4104. The substance was naive, but the semantics and logic were pre-structured with tagged segments and condition-action form, which is faster for a human to refine into a schema-conformant expression, than the original text.

This experiment suggests the potential to train a dedicated free/libre/open GPT (see: https://nlpcloud.com/gpt-3-open-source-alternatives-gpt-j-gpt-neo.html), or equivalent, on the body of the 40 FINRA rules which are known to be correctly tagged with the full taxonomy. The next 5 or 10 rules can be auto-transcribed, edited, and added to the reference set. After several iterations, the quality may become adequate to auto-transcribe the remaining 700+ rules accurately and consistently. Upon review, error patterns can be resolved, and the entire set can be re-processed.

7 OpenAI/GPT-3 https://platform.openai.com/playground

[D.3] Comment on proliferation vs harmonization of regulatory taxonomies.

"The extensibility of XML ensured its adaptability to any context and any level of granularity. At first this seemed optimal, but through this decade XML schema proliferation resulted in a complicated labyrinth of competing standards. Ironically, this undermined computational simplicity and speed. Domain-specific XML schemas seem suitable when a browser has a limited job to do in attaching semantics to displayed content of an individual site. But the large number and diversity of XML schemas which had come to be designed and implemented 'bottom-up' by diverse communities led to redundancy and inconsistency for the Semantic Web as a whole." Source: (Potvin, 2023, p. 134)

FINRA could take the lead in preparing and putting online a look-up table that maps the several schemas related the securities industry, including but not necessarily limited to the following:

- FINRA's "Machine-Readable Rulebook" taxonomy;
- FINRA's "Web Electronic Filing Transfer" (EFT) schema;
  - https://www.finra.org/filing-reporting/web-crd/web-eft-schema-documentation-and-schema-files
- The SEC's "Electronic Data Gathering, Analysis, and Retrieval" (EDGAR) schema;
  - https://www.sec.gov/info/edgar/specifications/form13fxmltechspec
- The Enterprise Data Management Council's and Object Management Group's "Financial Industry Business Ontology" (FIBO);
  - https://spec.edmcouncil.org/fibo/
  - https://www.omg.org/hot-topics/finance.htm

  "Financial Regulation Ontologies" (FRO);
  - https://finregont.com/xbrl/

[D.4] How could this initiative benefit from open-source collaboration?

This question is addressed in detail on our website, therefore we direct the reader to that page rather than copying that text verbatim: https://xalgorithms.org/organization#value

We note that in recent years FINRA has strategically embraced free/libre/open source software licensing and methods, both as user and contributor:
https://www.finra.org/about/technology
https://github.com/orgs/FINRAOS/repositories

3. Constraints of FINRA's Current 'Machine-Readable' Rulebook, and Ways to Transcend Them

What apparent constraints might prevent FINRA's Machine-Readable Rulebook initiative from meeting the requirements of its members' diverse conformance management approaches, and of its own market surveillance?

The stated purpose of FINRA's "Machine-Readable Rulebook" is "to enhance firms' compliance efforts, reduce costs and aid in risk management". In our assessment there are three specific constraints inherent in FINRA's approach as currently described and prototyped. Left unresolved, these issues could prevent the initiative from accomplishing the objective of improved conformance management among industry members, and of enhancing its own market surveillance systems.

3.1 "Natural Language from Rule-Makers" Versus "Natural Language for Rule-Takers"

Recommendation

> *FINRA's "Machine-Readable Rulebook" initiative is designed to facilitate finding the regulations that securities brokers and dealers must conform with. Future work could include a systematic approach to providing auxiliary natural language summaries that would enhance their situational recall and understanding*.

Rationale

Rule makers who draft legislation, standards, interpretations and guidelines are, of course, obligated to express themselves with precision. They need to refine the wording of each rule to ensure that it states exactly what is intended. On the other hand, rule takers intent on rules conformance must perform complex situational recall of numerous obligations/exclusions, permissions/prohibitions, and encouragements/discouragements. Although securities dealers and brokers typically hold university degrees in finance, accounting, economics or business, and prepare for and pass exams to obtain and maintain their licenses, even the most intelligent and honest among them face the "precision-recall tradeoff" described half century ago by Cyril Cleverdon:

> "As a general rule it remains true that in a large number of situations, an improvement in recall can only be obtained with a loss in precision, or vice versa, and it is reasonable to operate a system using this as a working principle. However, the inverse relationship of recall and precision is not a fundamental law..." (Cleverdon, 1972, p. 195, 199)

Simplification for sophisticated professionals requires choosing terms and phrasing optimized for understanding and recall of the essentials. A commonly known illustration is the 200-word summary of the 2,500-word *Creative Commons Attribution 4.0 International License.* It is introduced with the caveat: "This is a human-readable summary of (and not a substitute for) the license." [8]

Excerpt from the 'Data with Direction Specification'

"The Internet Engineering Task force specifies: " 'Simplification of language' here refers to ways of controlling expressions in a language to make reading or comprehension easier for particular target audiences". (Phillips & Davis, 2009) In the 1950s the UK Government had Ernest Gowers provide guidance in *Plain Words* for how to achieve straightforward communication, as this is indispensable to getting practical work done:

> "But what is this job that must be got on with? ... the writer's job is to make his reader apprehend his meaning readily and precisely. ... Even when he knows what he means, and says it in a way that is clear to him, is it always equally clear to his reader? If not, he has not been getting on with the job." (Gowers, 1954, p. 78)"

It was aeronautical engineer Clarence Johnson who emphasized "applying the simplest, most straight-forward methods possible to develop and produce new products" and then articulated the famous aphorism: "Keep it simple, stupid—KISS" (Rich, 1995, p 221, 231). System procedures, interfaces, and documentation, can benefit from the well-known 7±2 guideline that average human short-term memory capacity for processing information is constrained to about seven plus or minus two items (Miller, 1994), or its less prominent 4±1 refinement (Cowan, 2001) (Mathy & Feldman, 2012)." (Potvin 2023, p. 158, 75)

8 https://creativecommons.org/licenses/by/4.0/

3.2 "Machine Readable" Versus "Machine Processable"

Recommendation

> *FINRA's "Machine-Readable Rulebook" initiative is premised on use cases where computing resources can support highly-expressive semantic data processing. Future work could include specialized support for speed-optimized, in-memory key-value sifting methods suited to algorithmic high-frequency transaction systems.*

Rationale

High speed, high volume data processing at the scale performed by algorithmic, electronic, automated and high-frequency trading systems, and by FINRA's market surveillance systems, need to validate rule conformance without being slowed down by compute-intensive parsing of expressive sentences or hierarchies of semantic tags.

Applying a meaningful taxonomy to natural language data is suitable for use cases within conventional "Semantic Web" scenarios in which local browsers or interactive apps have the small job to do of associating meaning with displayed texts while interacting with a human. But this method of semantic tagging of expressive natural language is not usable for extremely high-speed high-volume normative data processing. Even among some of the most advanced methods of interactive natural language, complex semantics have been replaced with brute-force stochastics. (Vaswani et al., 2017)

Excerpts from the 'Data with Direction Specification'

There are numerous techniques available for optimizing a rule system for speed and throughput. Following are section headings that identify various techniques employed in that particular design:

5.3 Methods for High Performance Decentralized Distributed Computing
- 5.3.1 Externalize Computational Work from Run-Time
- 5.3.2 Externalize Complexity from Expression with Simple Controlled Natural Language
- 5.3.3 Externalize Linguistic Complexity from Rule Structure, to Simplify Function
- 5.3.4 Externalize Engagement of Semantic Web Standards to Rule Makers and Rule Takers
- 5.3.5 Externalize Computability by Requiring Rule Expression to be NOT Turing-Complete
- 5.3.6 Externalize Control Data and Logical Relations Data by Separating Data from Procedure
- 5.3.7 Externalize the Data Processing Burden with Purposeful Structuring of Data Into Tables
- 5.3.8 Externalize Reusable Algorithms (In-Memory Retrieval of Cartesian Product Tables)
- 5.3.9 Externalize Declarative Conditions and Assertions from Logical Relations

(Potvin 2023, p. 155-200)

"[O]ptimal' rule systems ... enable individuals and entities to communicate normative propositions more *cost-efficiently* and *cost-effectively* than is otherwise currently feasible:

- *Cost Effectiveness:* Maximize the quality of direction-intrinsic data communication within a given amount of time, resources and risk.
- *Cost Efficiency:* Minimize the time, resources and risk needed to achieve an intended quality of direction-intrinsic data communication." (Potvin 2023, p. 102)

3.3 “Rule Book” Versus “Rule System”

Recommendation

> *FINRA’s “Machine-Readable Rulebook” initiative involves delivery of two online services: Web-based rules search, and an API for rulesbase queries. Future work could include free/libre/open collaborative experimentation with end-to-end systems to advance the normative performance of the US securities market.*

Rationale

The statistics in the "Regulatory Actions and Corporate Financing Review 2017–2021" online at https://www.finra.org/media-center/statistics are worth some reflection. In those five years the number of investor complaints received by FINRA has nearly quintupled, and yet the number of disciplinary actions filed, and the number of individuals barred and suspended, each declined by almost half. No interpretation of these apparently contradictory trends (greater rules conformance, yet lower investor protection?) is provided in FINRA’s 2021 Annual Report.[9] Perhaps there were fewer violations of the rules overall, but those transgressions which did occur affected many more investors, more severely.

A whole systems perspective on FINRA’s “Machine-Readable Rulebook” initiative considers the general trends and dynamic forces shaping rules communication, surveillance, response and outcome. As a dynamic interactive phenomenon, FINRA inevitably faces *"The Problem-Solvers’ Paradox":* the greater and more sustained FINRA’s success in terms of rules conformance, the lower the perceived need for its services, which can weaken vigilance and increase vulnerability to fewer but more severe abuses. A systems designer considers ways to re-frame this dynamic, for example one might brainstorm a *Market Integrity Index Fund* that would increase in value as verifiable normative performance indicators demonstrate improvements in both rules conformance and investor protection.

Excerpt from the ‘Data with Direction Specification’

This design research provides a rationale, a functional specification and partial prototype working components to solve the following general class of problem:

> *Agent A, interacting with Agent B, requires knowledge of one or more externally-managed rules from Agents C..n that are ‘in effect’ for given contexts, and are ‘applicable’ to a set of event categories, and are ‘invoked’ by particular circumstances, where:*
>
> *(i) A and B may or may not know about C..n’s rules, or about any updates to them, but either or both would prefer to obtain all available facts about relevant rules when interacting.*
>
> *(ii) C..n may or may not know about A and B in particular, nor about their particular medium of interaction, but can expect A or B or their medium of interaction to be capable of exchanging data with a generic medium common to A..n.*
>
> *(iii) A and B would tolerate the risk of exposing limited data through the generic medium so that it can be used to select information about relevant rules from C..n.*

...The “Data With Direction Specification” (DWDS) describes a type of distributed, general purpose system that individuals and organizations can use to author, publish, discover, fetch, scrutinize, prioritize and, with agreement of direct stakeholders, automate rules across any informatics network with precision, simplicity, scale, speed, resilience, and deference to prerogative. DWDS describes a class of data-processing pipeline with the underlying relation: 'IS + RULE ⟹ OUGHT'. (Potvin 2023, p. 57, 146)

9 https://www.finra.org/sites/default/files/2022-06/2021-FINRA-Financial-Annual-Report.pdf

4. Our Perspective on the Potential of FINRA's 'Machine-Readable' Rulebook

> ***How might FINRA's Machine-Readable Rulebook be adapted to improve human comprehension and recall of the rules; to meet the speed and volume requirements of algorithmic transactions; and to reduce the rules management burden?***

The Xalgorithms community perspective on rule systems design is detailed in a recently-completed 250-page thesis, which we include as supporting documentation to the present submission. The GitLab URL provided below supplies the most recently edited version, and an overview presentation deck.

Potvin, J. (2023). *Data with Direction: Design Research Leading to a System Specification for 'an Internet of Rules'. Dissertation in partial fulfillment of a Doctorate in Business Administration—DBA (Project Management).* Université du Québec—Outaouais Campus (UQO)]. *License CC-by 4.0*
https://gitlab.com/xalgorithms-alliance/data-with-direction-specification/dwds-documents/-/tree/master/current

The following three sections highlight elements of how FINRA's "Machine-Readable Rulebook" can be integrated with the DWDS "Internet of Rules" concept and functional design to advance conformance management and investor protection through improved human access to, as well as comprehension and recall of the rules; to facilitate high performance operationalization of FINRA's rules in algorithmic transaction and surveillance systems; and to reduce FINRA's internal rules maintenance workload.

4.1 Situating FINRA's Rulebook in the DWDS 'Internet of Rules' System Concepts and Functions

Figure 1: A View of FINRA, its Rulebook, and its Members in the Conceptual Space of an End-to-End Rules System. Adapted from: (Potvin 2023, Fig. 10, p. 148)

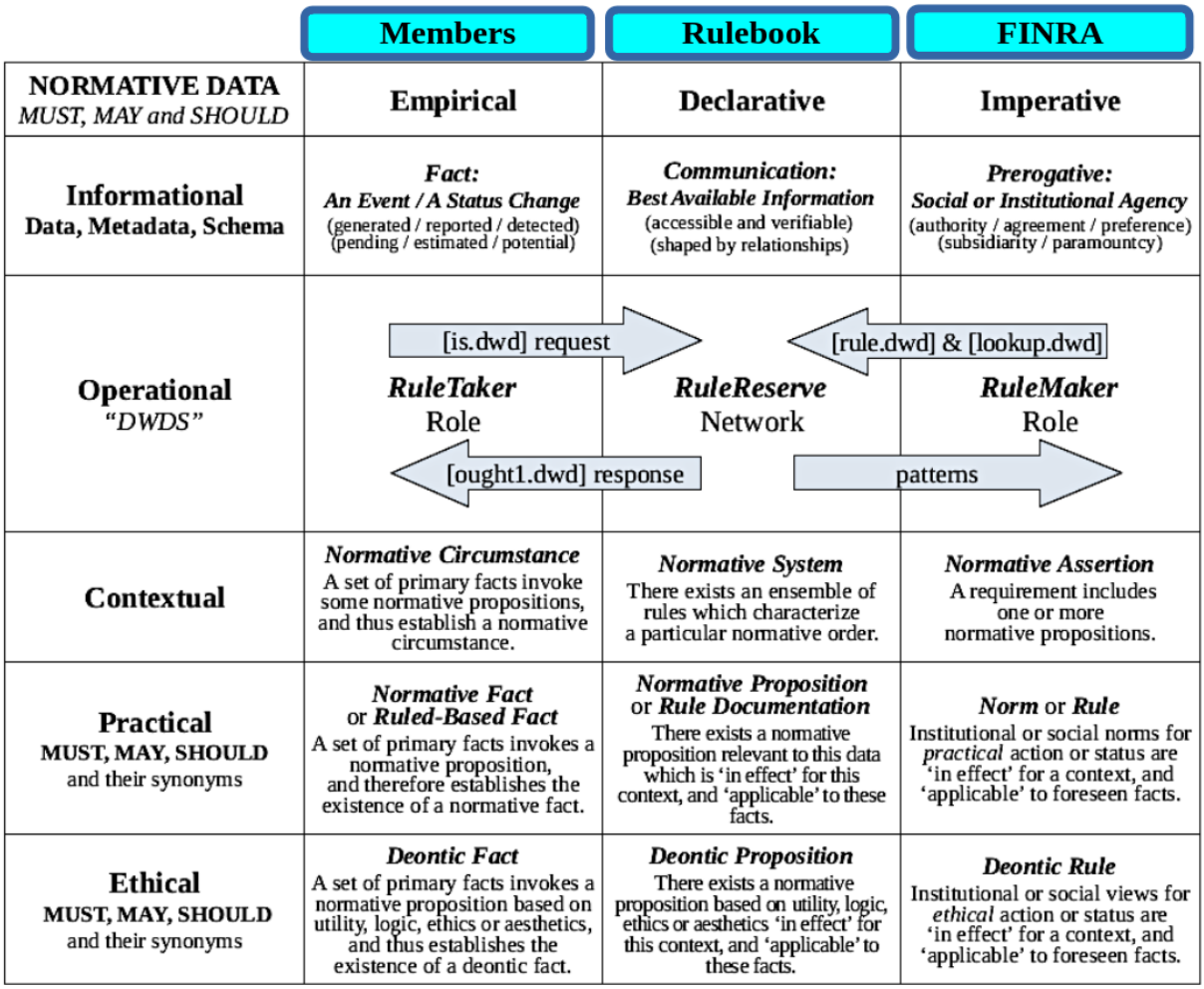


Figure 2: The Functional Role of FINRA in the "Rule Maker Role" of the DWDS Sequence Diagram. Adapted from: (Potvin 2023, Fig. 11, p. 150)

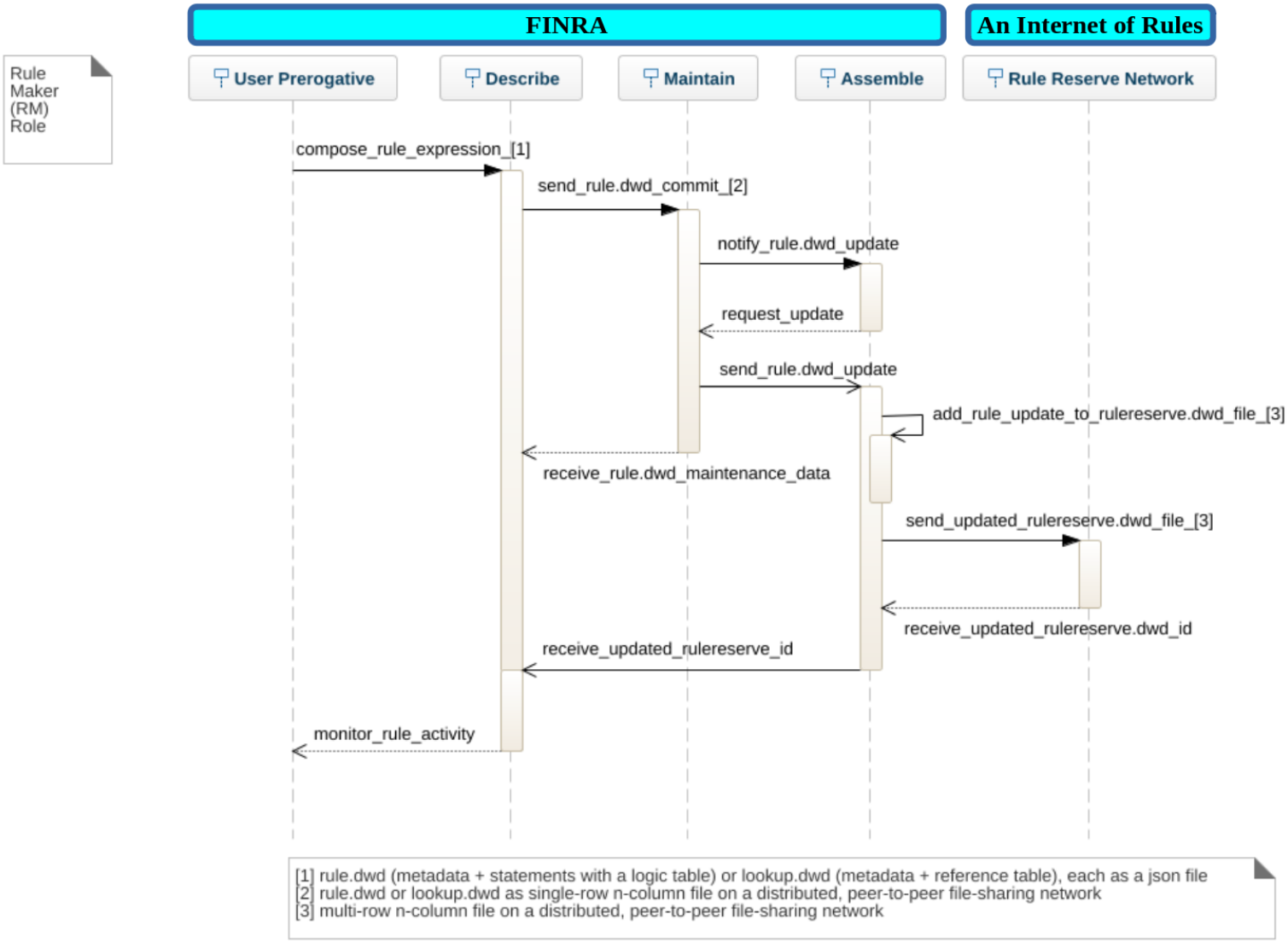


Figure 3: The Functional Role of the FINRA's Rulebook in the "Subset Rule Reserve Role" of the DWDS Sequence Diagram. Note that this only shows the top half of the Rule Reserve Network functions. Adapted from: (Potvin 2023, Fig. 12, p. 151)

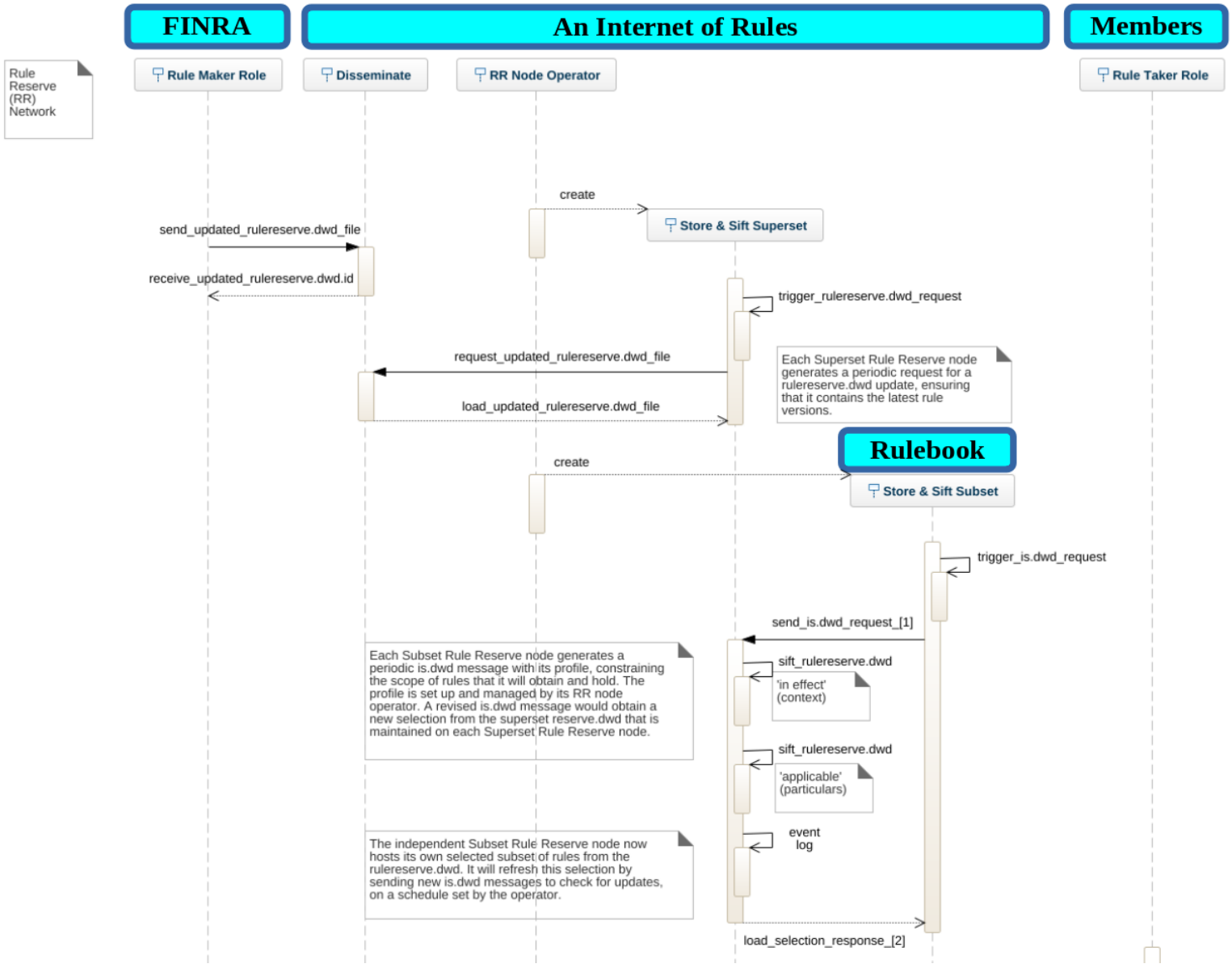
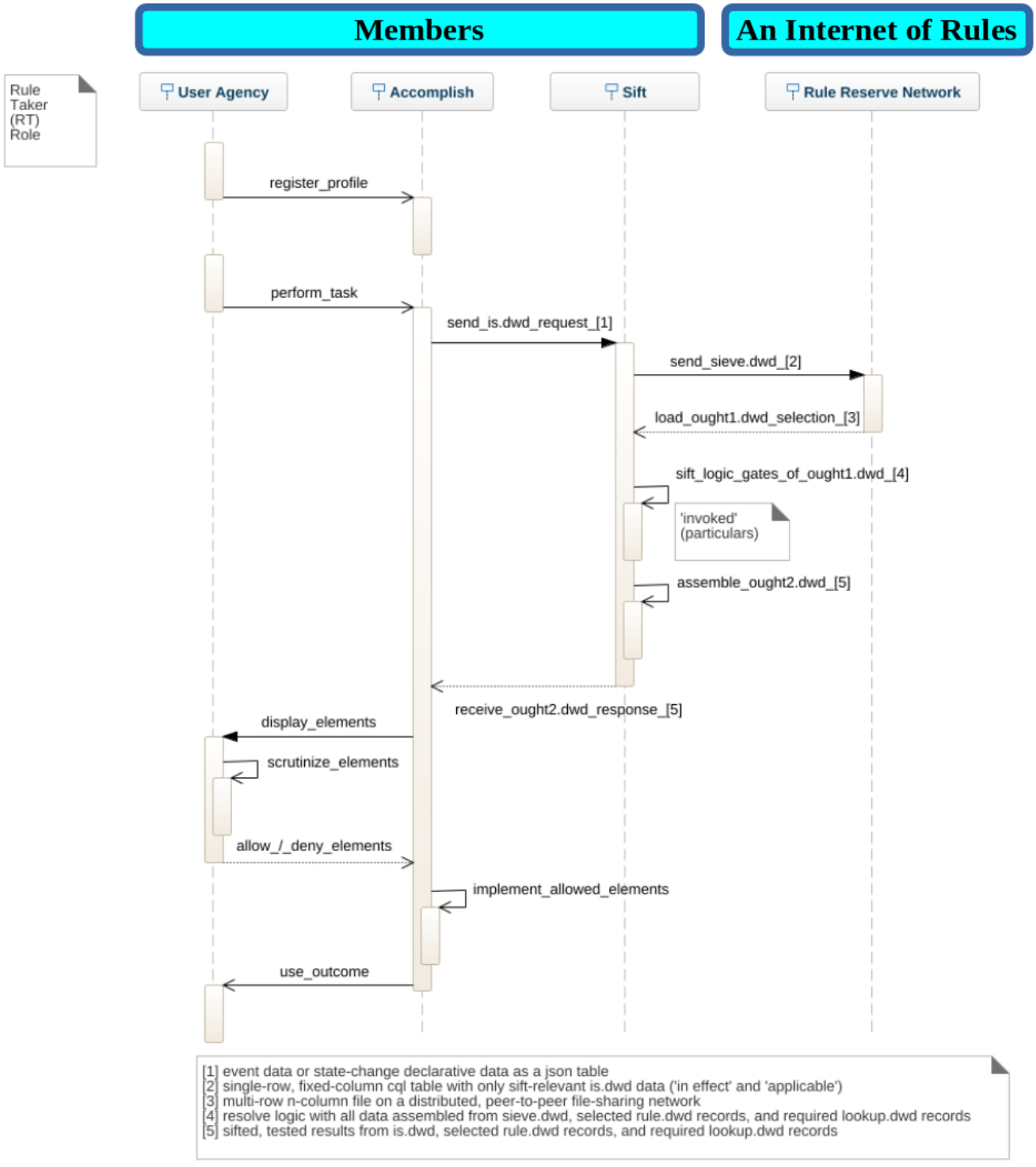


Figure 4: The Functional Role of FINRA Member Firms and Individuals in the "Rule Taker Role" of DWDS Sequence Diagram. Adapted from: (Potvin 2023, Fig. 13, p. 152



4.2 Transcribing Two Sample FINRA Rules to DWDS RuleData Form

For this submission two sample FINRA rules (3240; 4140) were structured into DWDS RuleData form. The corresponding JSON files are attached as supporting documentation, and working views of their respective logic gates are provided on the following pages, as portrayed in the graphical user interface of the RuleMaker Web app. The images are also attached separately to facilitate reading of the small text of the sentences that comprise the Input Conditions and Output Assertions.

The reader should take into account the following considerations:

(a) The permutation scenarios are to be read vertically, as Scenario A, Scenario B, and so on.

(b) The symbols have the following meanings:

Figure 5: Meanings of Symbols in DWDS Logic Gates.
Adapted from (Potvin 2013, Fig 19, p. 192)

Input Symbol	Input Conditions	Output Symbol	Output Assertions
X	NO	⊘	NOT
✓	YES	!!	MUST
&	YES AND NO	○	MAY
?	YES OR NO	!	SHOULD

(c) The sentences of the Input Conditions and Output Assertions were adapted in three ways:

i. As discussed in Section 3.1 of this submission, the natural language of rule-makers is not necessarily the same as natural language suited to rule-takers. The metadata for each rule include the URL to the original regulatory text so users can readily consult the original.

ii. Each sentence is ‘shoehorned’ into DWDS “finite state grammar” of 6 syntactic elements;

iii. Each sentence has been adapted to the RuleSpeak guidelines to the extent practical within the constraints of the DWDS six syntactic elements (e.g. Use simple language; Break rules into atomic sentences; Avoid ambiguities; etc…). (Ross, 2023) (OMG, 2016)

iv. To experiment with a potential time-saving method, I instructed a vanilla (GPT)[10] as follows:

```
Re-write the following rule using only discrete declarative
sentences in a style that conforms with the essential
practices of "RuleSpeak", starting a new line for each
sentence, and without leaving out any operational details or
references. "4140. Audit. (a) FINRA may at any time...
```

This successfully transformed the original text of the sample regulation into well-structured declarative sentences. To further expedite the process of inserting the sentences into the six syntactic elements of DWDS RuleData I instructed the GTP with this:

```
Identify the 'subject', the 'predicate', and the 'object' in
each of the following sentences...
```

This saved some time, but with mediocre results. Probably a GTP could be trained with a set of declarative sentences pre-partitioned as described, to obtain higher accuracy.

10 OpenAI/GPT-3 https://platform.openai.com/playground

This representation of the logic structure of a rule in the DWDS RuleMaker Web app facilitates discussion and refinement of the individual sentences for the Input Conditions and Output Assertions, and well-organized consideration of the potential permutations to be anticipated in the logic relations.

Figure 6: Version 0.1.0 of a DWDS Logic Gate for FINRA Rule 3240: "Borrowing From or Lending to Customers", as seen in the RuleMaker interface.

Figure 7: Version 0.1.0 of a DWDS Logic Gate for FINRA Rule 4140: "Audit", as seen in the RuleMaker interface.

The RuleMaker working environment provides for both "machine readable" and "machine processable" rule expression. The JSON record of this logic gate, which includes the rule metadata and optional descriptive fields, is auto-generated by the RuleMaker Web app. It can saved to one's local drive and/or published to the Internet (IPFS) on any node of the RuleReserve Network.

This entire rules management process is in the hands of subject matter experts, while software programmers are focused on ensuring that the enabling applications are working properly. There is no requirement for software programmers to interpret regulation semantics or the rule logic.

In the examples provided here in Figures 6 and 7, the sentences have not yet been aligned to FINRA's semantic taxonomy. That work requires more familiarity with the particular controlled natural language schema than the present author currently possesses. However this version 0.1.0 provides a convenient venue for collaboration to do so. FINRA's taxonomy would be applied to the sentence elements. Discussion is required to determine exactly how this should be performed and displayed in RuleMaker.

Once all the metadata, logic gate data, and optional descriptive data are entered into the RuleMaker Web app for a rule, the user can have it automatically generate the JSON file for local storage and/or Internet publication to IPFS, which supplies a unique Content Identifier (CID) for that precise version of that rule. Below is part of the JSON record for Rule 4140, from Figure 7.


```
{
    "id": "ce4c3fa7-1c84-4f00-8cc5-dfc11eee947c",
    "uuid": "ce4c3fa7-1c84-4f00-8cc5-dfc11eee947c",
    "rule_id": "ce4c3fa7-1c84-4f00-8cc5-dfc11eee947c",
    "rulereserve_nodes": "*",
    "version_standard_url": "https://semver.org/",
    "dwds_schema_version": "0.0.0",
    "properties": {
        "id": "ce4c3fa7-1c84-4f00-8cc5-dfc11eee947c"
    },
    "metadata": {
        "rule": {
            "120_title": "4140. Audit",
                "240_summary": "FINRA Rules\n4000. FINANCIAL AND OPERATIONAL RULES\n4100. FINANCIAL CONDITION\n4140. Audit",
            "960_explanation": "FINRA may require any member to have an audit or examination of its accounts conducted by an independent public accountant. The audit or examination must follow attestation, review, and consultation standards specified by the AICPA and any additional requirements set by FINRA. The audit or examination is directed by FINRA's Executive Vice President in charge of financial responsibility, or a delegate of theirs. Any member who does not file the relevant audited financial and/or operational report or examination report within the given timeframe will be subject to a late fee listed in Schedule A Section 4(g)(1) of the FINRA By-Laws.",
            "version": "0.1.0",
            "criticality": "experimental",
            "url": "https://www.finra.org/rules-guidance/rulebooks/finra-rules/4140",
            "pattern": "",
            "pattern_version": "",
            "rulemaker_entity": [
                {
                    "name": "Financial Industry Regulatory Authority (FINRA)",
                    "url": "https://www.finra.org/",
                    "uuid": "027d7f9e-0e7d-44cd-a893-5c12c6a20d0b"
                }
            ],
            "rulemaker_manager": [
                {
                    "name": "Xxxx",
                    "email": "Xxxx@finra.org",
                    "contact": "General Counsel, Office of the General Counsel",
                    "uuid": "b88e561c-2d5c-464b-a88a-de66f45096f0"
                }
            ],
            "rulemaker_author": [
                {
                    "name": "Joseph Potvin",
                    "email": "jpotvin@xalgorithms.org",
                    "contact": "",
                    "uuid": "49354a2a-fad3-44f8-bb0d-3891e6ed0d34"
                }
            ],
            "rulemaker_maintainer": [
                {
                    "name": "Joseph Potvin",
                    "email": "jpotvin@xalgorithms.org",
                    "contact": "",
                    "uuid": "4f2e65f7-e22e-40a7-a705-b3e89c8226fb"
                }
            ]
        }
    },
    "in_effect": [
        {
            "country": "US",
            "subcountry": "",
            "timezone": {
                "start": "UTC-05:00",
                "end": "UTC-05:00"
            },
            "start": "1992-08-12T04:00:01.000Z",
            "end": "2011-08-12T04:59:59.000Z"
        }
    ],
```

```
    "category_applicable": {
        "industry": [
            {
                "isic_code": "6611",
                "isic_name": "Administration of financial markets"
            }
        ],
        "good_service_asset": [
            {
                "unspsc_code": "64110000",
                "unspsc_name": "Securities"
            }
        ]
    },
    "data_sources": [],
    "input_conditions": [
        {
            "sentence": [
                {
                    "determiner": "This"
                },
                {
                    "noun": "entity"
                },
                {
                    "predicate_verb": "is"
                },
                {
                    "description": "a registered member"
                },
                {
                    "attribute": "of FINRA (FINRA member entity),"
                },
                {
                    "past_participle_verb": "as validated in the FINRA membership registry."
                }
            ],
            "scenarios": {
                "A": "01",
                "B": "01",
                "C": "01",
                "D": "01",
                "E": "01",
                "F": "01",
                "G": "01",
                "H": "01"
            }
        },
        {
            "sentence": [
                {
                    "determiner": "This"
                },
                {
                    "noun": "FINRA member entity"
                },
                {
                    "predicate_verb": "has been instructed to file"
                },
                {
                    "attribute": "within a given time frame,"
                },
                {
                    "description": "an audited financial and/or operational report or
examination report to validate the accuracy or integrity of its financial statements, books
and records or prior audited financial statements,"
                },
                {
                    "past_participle_verb": "as instructed."
                }
            ],
            "scenarios": {
                "A": "00",
                "B": "01",
                "C": "01",
                "D": "01",
                "E": "01",
                "F": "01",
                "G": "01",
                "H": "01"
            }
        },  ...
```


(The JSON representation of the logic gate continues, followed by optional descriptive data.)

4.3 Draft Charter for a Financial Securities Regulations Working Group

In Section 1 of this submission we explained that Working Groups hosted by Xalgorithms Foundation have their own written charter, managing their own donated funds. Participants can include businesses, governments, academics and not-for-profits, civil society communities and individuals, collaborating under Xalgorithm's 100% free/libre/open source model based on the Apache 2.0 license (RuleMaker Web app and RuleTaker embedded component) and the AGPL 3.0 license (RuleReserve network service), while documentation and data are shared under the CC-by 4.0 International license. One or more "Contributor Agreements" can be appended to a charter, and tailored to circumstance.

Following is 'first draft' working text towards the potential charter for a "Financial Securities Regulations Working Group".

Draft for Discussion

Issue to be Addressed: Financial Securities Market Integrity and Investor Protection

A well-functioning financial securities market operates on sets of rules and a cost-effective, cost-efficient generic rules system. Market integrity depends on human accessibility, comprehension and recall of those rules, and on high performance operationalization of the rules in algorithmic transaction and surveillance systems.

Requirement: On-Demand Delivery of Regulations 'In-Effect', 'Applicable' and 'Invoked'

Financial securities dealers and regulators have a common interest in event-triggered transmission of concise, current, and correct information about normative rules that are: 'in effect' for given dates/times, identities and jurisdictions; 'applicable' to a set of industry and product/service categories; and, 'invoked' by particular event circumstances; in a manner that is readily comprehensible to humans and directly usable in high-performance applications and platforms.

The behavioural and operational aspects of financial securities regulations are far more likely to be understood and conformed with when simple human-readable and fast machine-processable assertions of MUST, MAY and SHOULD (or their syonymns or negatives) are delivered on-demand to individuals, organizations and/or their machines at the instant they are relevant.

Proposed Approach: The Data With Direction Specification (DWDS) for "an Internet of Rules"

The "Data With Direction Specification" (DWDS) operationalizes the essential conceptual relation:

'IS + RULE ⟹ OUGHT'

The specification describes a type of distributed, decentralized, general purpose end-to-end data-processing pipeline that individuals and organizations can use to author, publish, discover, fetch, scrutinize, prioritize and, with agreement of direct stakeholders, automate rules across any informatics network with precision, simplicity, scale, speed, resilience, and deference to prerogatives, agreements and preferences. The functional design involves a RuleData data structure suitable for any platform and any language, a RuleMaker application with the *imperative* role in normative communication (i.e issuing rules), a RuleReserve network service with the *declarative* role (identifying rules that are 'in effect' for a context and 'applicable' to a set of categories), and a RuleTaker component with the *empirical* role (sending a set of circumstantial facts and receiving facts about rules deemed to be invoked by the those facts). Operated together these give rise to an "Internet of Rules" – a method by which independent, self-contained rules are transmitted efficiently and flexibly from the source repositories in which they are maintained, to the applications that use them.

<u>One-Year Workplan</u>: April 2023 to March 2024

Following is a tentative one-year project schedule oriented to the delivery of interim results and management check-points. This serves as a guide only, to be updated as determined by participants.

- **Month 1:**
  - Use the present one-year Working Group plan to elaborate particular objectives for various community contributors, and to frame the relationships with stakeholders.
  - Create an effective participatory R&D collaborative trajectory involving participants from multiple data supply organizations.
  - Identify a sample of rules for community testing, that range from simple to complicated.
  - Design structured test protocols for RuleTaker implementations in at least three widely deployed production-class algorithmic trading systems currently in use for securities.
  - Adapt or create a basic online test service for validating automated rule conformance:
    - Multilingual, accessible (WCAG 2.0) end-user interface.
    - Rapid iterative diagnosis and documentation of discrepancies.
    - Comprehensive task management workflow.

- **Months 2-3-4:**
  - Test transaction scenaros with RuleMaker, RuleReserve, RuleTaker reference implementations.
  - Incrementally increase rule complexity; refining the process for accuracy and for speed.
  - Refine the online service for validating rule conformance.
  - Develop a draft risk management model of “Internet of Rules” users**.**
  - Jointly develop and present a first interim report to stakeholders.

- **Months 5-6-7:**
  - Incrementally broaden collaborative work on rule expression and validation.
  - Roll out and support version 1.0 of the online service.
  - Publish version 1.0 documentation (technical, financial, legal).
  - Broaden consultations (technical, financial, legal).
  - Create hypothetical management/financial models for proliferation.
  - Jointly develop and present a second interim report to stakeholders.

- **Months 8-9-10:**
  - Increase collaborative work on rule expression and validation.
  - Test and debug complicated rules, exceptions, anomalies and dependency chains (forward-chained, backward-chained rules).
  - Commence scheduled version updates for each quarter (3 months).
  - Refine and publish documentation (technical, financial, legal).
  - Jointly develop and present a third interim report to stakeholders.

- **Month 12**
  - Develop for discussion and refine a workplan for Year 2.
  - Contract out an arms-length study for stakeholder/community views.
  - Assess demand for training, and make arrangements accordingly.
  - Assess demand for support, and make arrangements accordingly.

Reference List

Cleverdon, C. W. (1972). On the Inverse Relationship of Recall and Precision. *Journal of Documentation*, *28*(3), 195–201. https://doi.org/10.1108/eb026538

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*(1), 87–114. https://doi.org/10.1017/S0140525X01003922

Gowers, E. (1954). *The Complete Plain Words*. UK Stationary Office.

Mathy, F., & Feldman, J. (2012). What's magic about magic numbers? Chunking and data compression in short-term memory. *Cognition*, *122*(3), 346–362. https://doi.org/10.1016/j.cognition.2011.11.003

Miller, G. A. (1994). The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information. *Psychological Review*, *101*(2), 345–352.

OMG. (2016). *The RuleSpeak Business Rule Notation. Semantics of Business Vocabulary and Business Rules (SBVR), v1.4*. Object Management Group. https://www.omg.org/cgi-bin/doc?formal/19-10-07.pdf

Phillips, A., & Davis, M. (2009). *Tags for Identifying Languages* (Request for Comments RFC 5646). Internet Engineering Task Force. https://doi.org/10.17487/RFC5646

Potvin, J. (2023). *Data with Direction: Design Research Leading to a System Specification for 'an Internet of Rules'. Submitted to the Board of Examiners in partial fulfillment of a Doctorate in Business Administration—DBA (Project Management). License CC-by 4.0* [Université du Québec—Outaouais Campus (UQO)]. https://gitlab.com/xalgorithms-alliance/data-with-direction-specification/dwds-documents/-/tree/master/current

Rich, B. (1995). *Clarence Leonard (Kelly) Johnson (1910-1990): A biographical memoir*. National Academy Press. http://www.nasonline.org/publications/biographical-memoirs/memoir-pdfs/johnson-clarence.pdf

Ross, R. (2023). *Rules: Shaping Behavior and Knowledge* (1st edition). Business Rule Solutions, LLC. https://www.brsolutions.com/rules-shaping-behavior-and-knowledge-book.html?src=BRC

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). *Attention Is All You Need* (arXiv:1706.03762). arXiv. http://arxiv.org/abs/1706.03762

Contact:

Dr. Joseph Potvin, DBA MPhil
Executive Director, Xalgorithms Foundation
50 Hines Road, Suite 240
Ottawa, Ontario, Canada K2K 2M5

(819) 593-5983
(57) 318-2000-933

jpotvin@xalgorithms.org
https://xalgorithms.org

Specification
https://gitlab.com/xalgorithms-alliance/data-with-direction-specification/dwds-documents/-/tree/master/current
License: CC-by International 4.0

Source Code
https://gitlab.com/xalgorithms-alliance
Licenses: Apache 2.0 (RuleMaker & RuleTaker); Affero GPL 3.0 (RuleReeserve)

Task Management
https://xalgorithms.redminepro.net

Notices
@Xalgorithms