AI4SE: Semi-Automated Development of Textual Requirements: Combined Natural Language Processing and Multi-Domain Semantic Approach

By
UMD Team: Mark Austin, Sachraa Borjigin, Edward Zontek-Carney
Early Contributors: Andres Arellano, Leonard Petnga
September 21-22. 2022

AI4SE/SE4AI Workshop

www.sercuarc.org

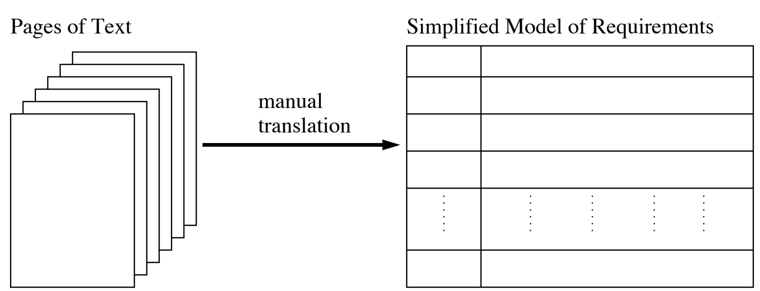
Cleared for Public Release

Motivation: Formal Modeling of Requirements

Modern Engineering Systems

Nearly always designed, built, operated and maintained by teams of people + automation. Key component in bringing this capability together is the ability to write and manage textual requirements early in the system development lifecycle.

State-of-the-Art Practice



Key Problems

- Poor writing skills; excessive use of acronyms.
- Manual translation process is slow and error-prone.
- How do we know if a system description is complete and consistent?

Motivation: Formal Modeling of Requirements

Recent Advances

- Increased adoption of SysML in the profession – semi-formal representations enhance quality of communication among team members, reduce uncertainties and improve project schedule.
- Remarkable advances in AI (semantic modeling and reasoning) and ML.
- Use of text persists (e.g., for standards, regulations, etc).

Formal Representation of Requirements

- Correct (syntax and semantics) encoding of textual requirements.
- Automatic processing and traceability in the design process.
- Specific, verifiable, realistic, time-bound.

Req: Within the ***school area***, ***maximum*** ***speed*** of vehicles shall be ***15*** mph ***when children are present.***

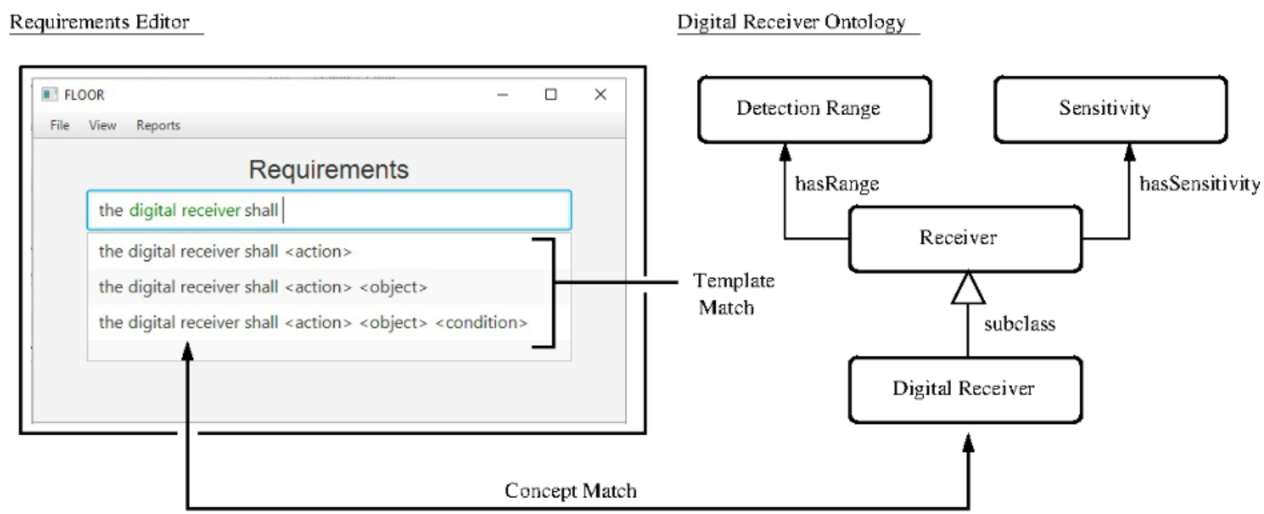
space | constraint | Physics | Phys. Qty | event

Motivation: Formal Modeling of Requirements

Brainstorming Project Ideas (Fall Semester, 2013)

Idea 1 (Syntax): Use knowledge in semantic graphs (ontologies and rules) to write template-compliant textual requirements.
Idea 2 (Semantics): Support writing (development) of semantically compliant requirements
Idea 3 (Integration): Provide automated support for verification of requirements at various levels of design.

So how might this work?

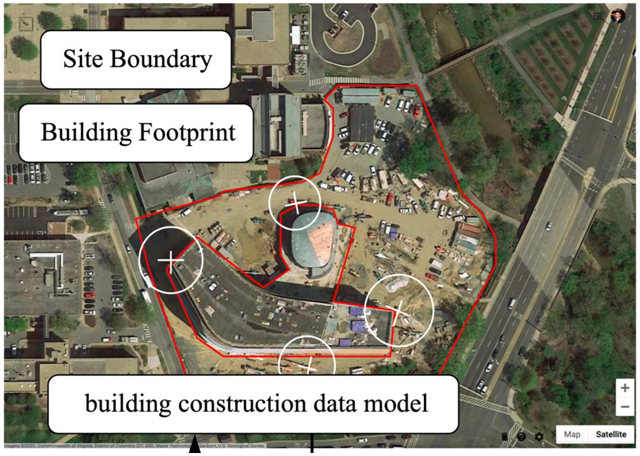


Presentation Objectives

Research Questions

- How can multi-domain semantic modeling and natural language processing work together to improve the development of textual requirements?
- How can requirements templates improve quality of validation of individual requirements and groups of requirements?

Motivating Case Study: New Computer Science Building at UMD



Sample Textual Requirements

Building Construction Requirements

Project Cost

1.1. The construction project budget shall be no more than US $150 million dollars.

Project Schedule

2.1. The construction project start date shall be on January 20, 2016.

2.2. The construction project end date shall be on April 20, 2019.

2.3. The construction project duration shall not exceed 4 years.

Building Foundation

3.1. The building foundation shall be constructed of solid materials.

3.2. The building foundation depth shall be at least 5 m.

3.3. The foundation wall thickness shall be no less than 15 cm.

3.4 The clear distance of building foundations to the construction site boundary shall be at least 60 cm.

Sample Textual Requirements

Building Construction Requirements

Construction Site

4.1. The construction manager is required to develop a project-specific orientation for all workers.

4.2. The general contractor shall have a process of validating training and certification in place for all construction workers.

Construction Equipment

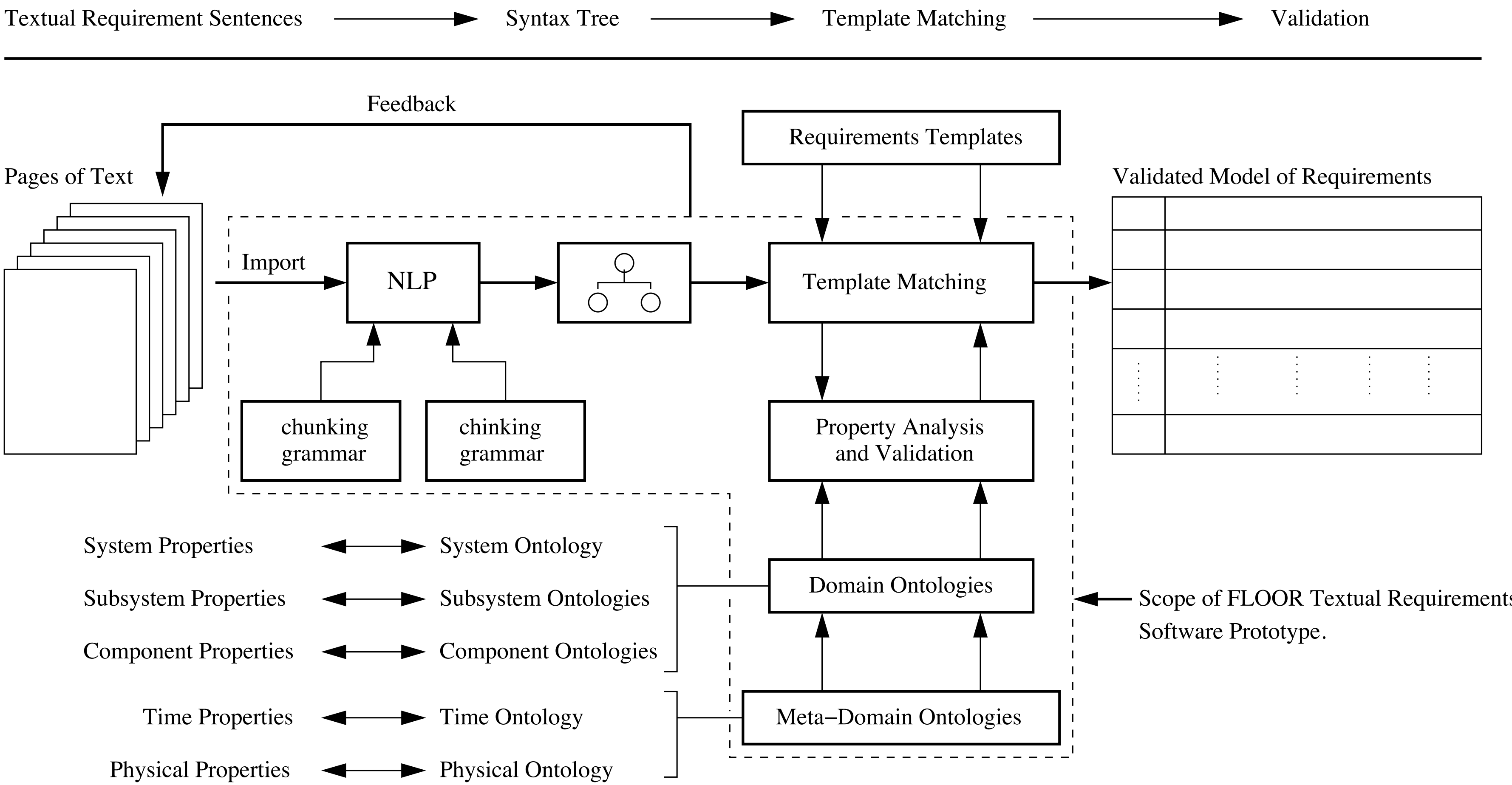
5.1. The general contractor shall maintain documentation of all equipment inspection.

5.2. The crane operations shall be suspended if the wind speed is greater than 40 km/hr.

Construction Workers

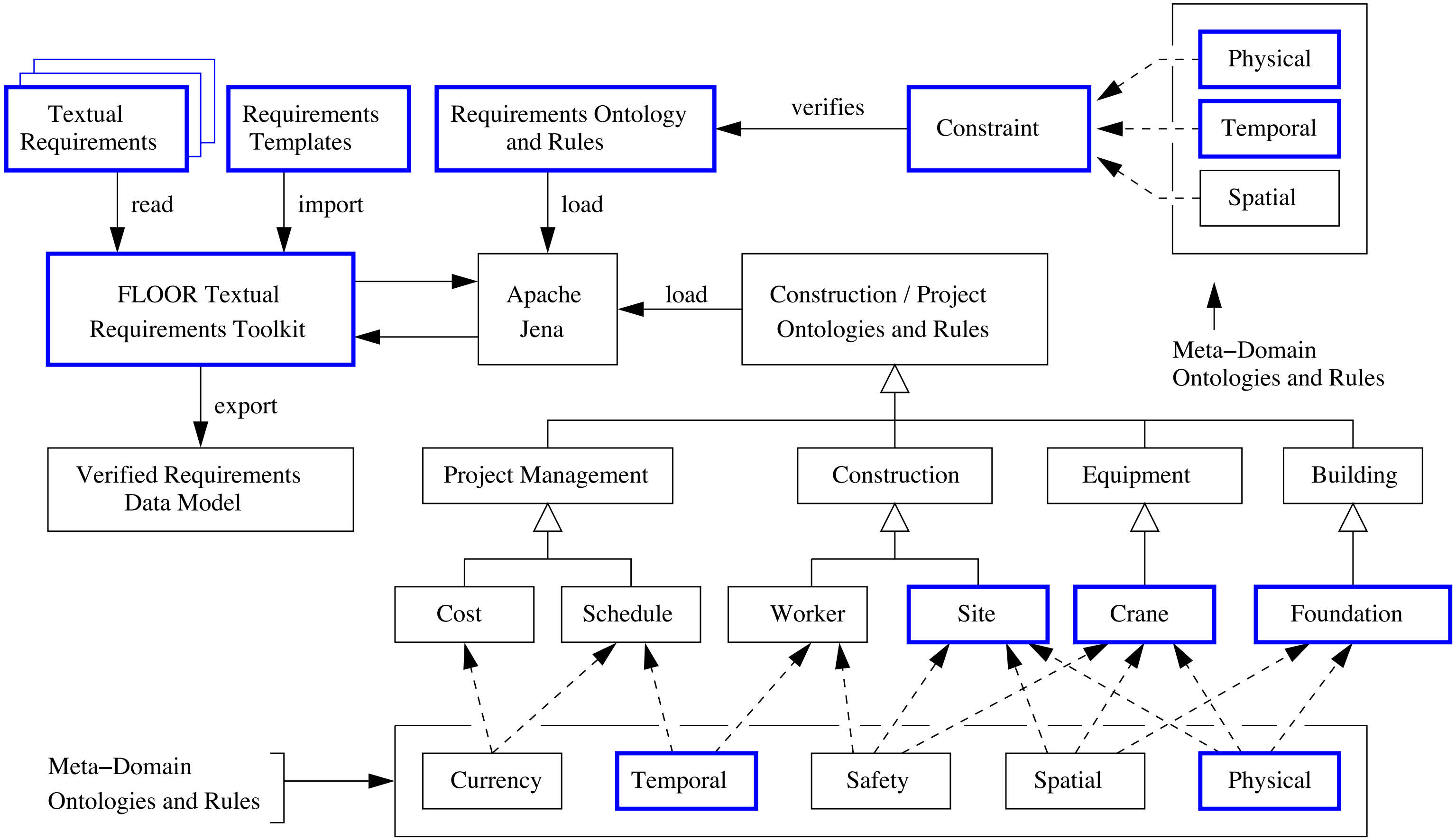
6.1. The construction workers shall remain a minimum of 6 m distance from any solid red labeled exhaust system in the construction site.

Proposed Approach (What's New?)



Source: Zontek-Carney, 2017.

Proposed Approach (What's New?)

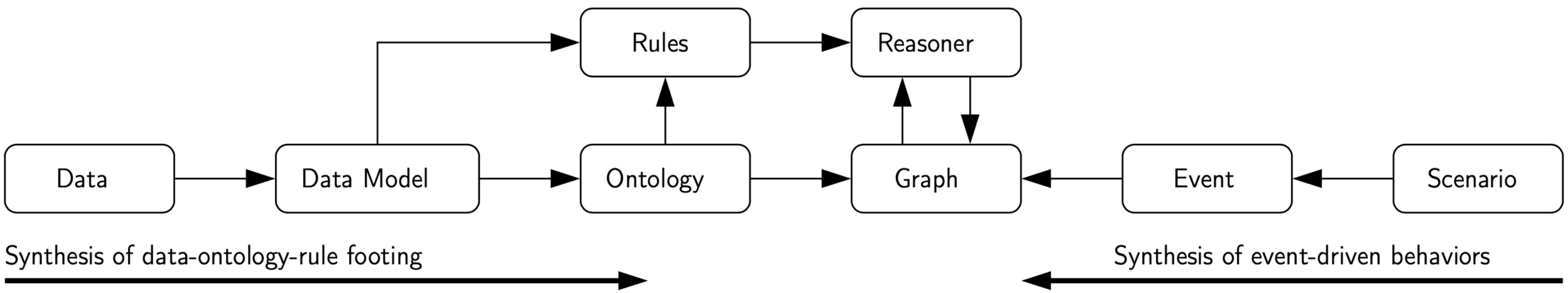


Legend for Graphical Notation: Round rectangles for classes and primitive data types.
Boxes represent software modules and individuals in semantic model.
Solid and dashed one–way arrows represent data property and object property relationships.

Source: Borjigin, Austin, and Zontek-Carney, 2022

Multi-Domain Semantic Modeling

Synthesis of System Behavior and System Structure

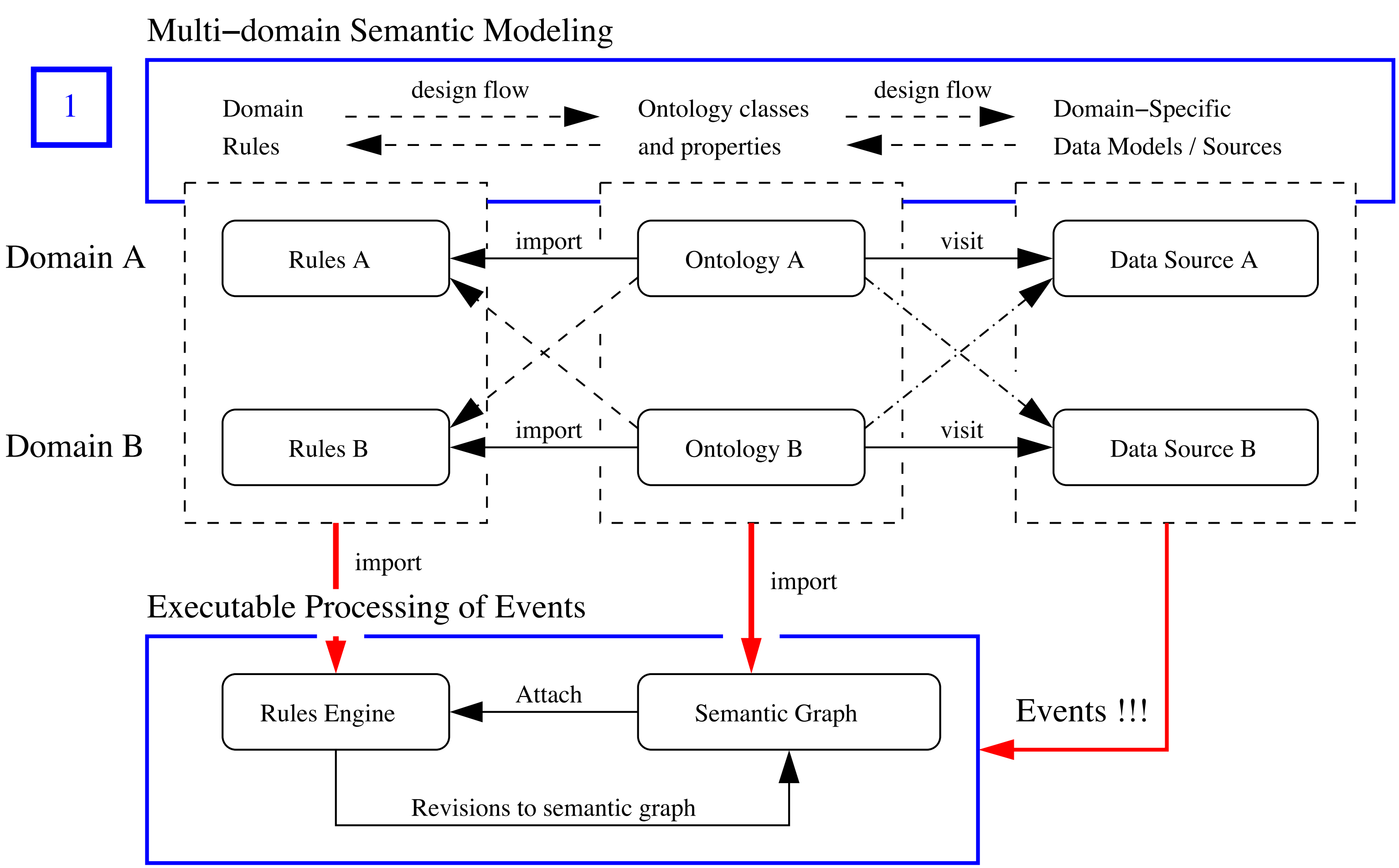


Guiding Principles

- Data-driven approach.
- Co-development of ontologies, rules and data models.
- Ontologies visit data models to get individuals.
- Enhance power of rules with backend software functions.
- Semantic graph dynamically responds to incoming events.

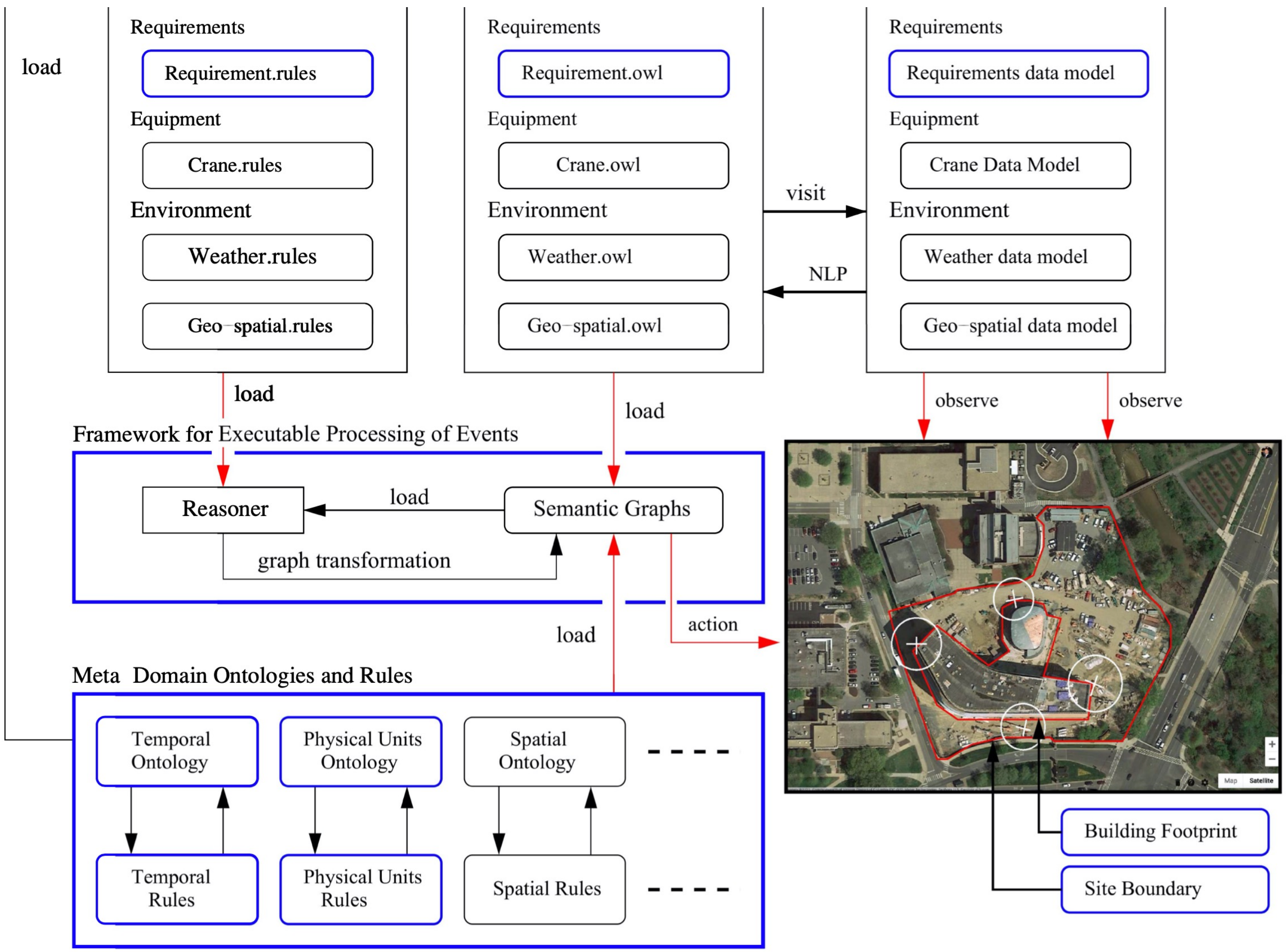
Multi-Domain Semantic Modeling

Data-Ontology-Rule Footing (UMD / NIST / SERC in 2017).



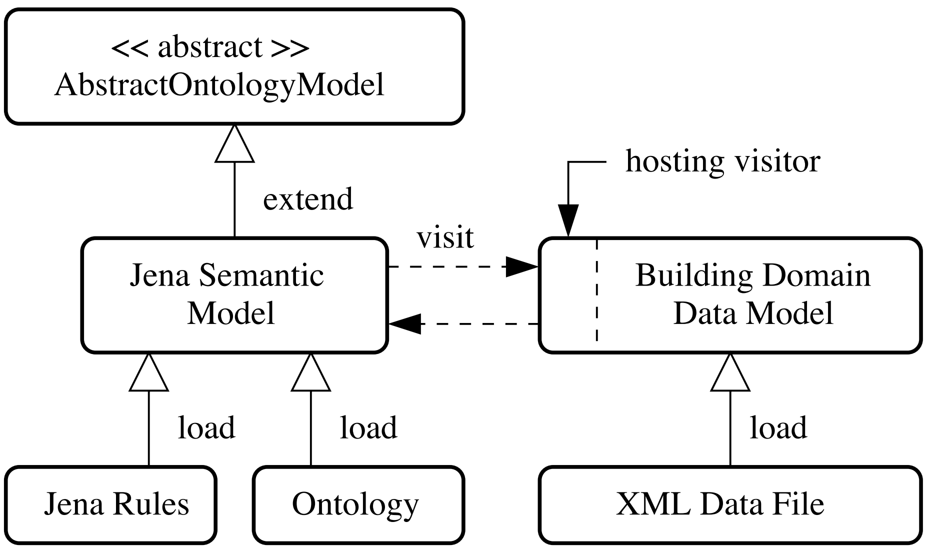
Source: Coelho, Austin, Blackburn, 2019

Executable Processing of Events

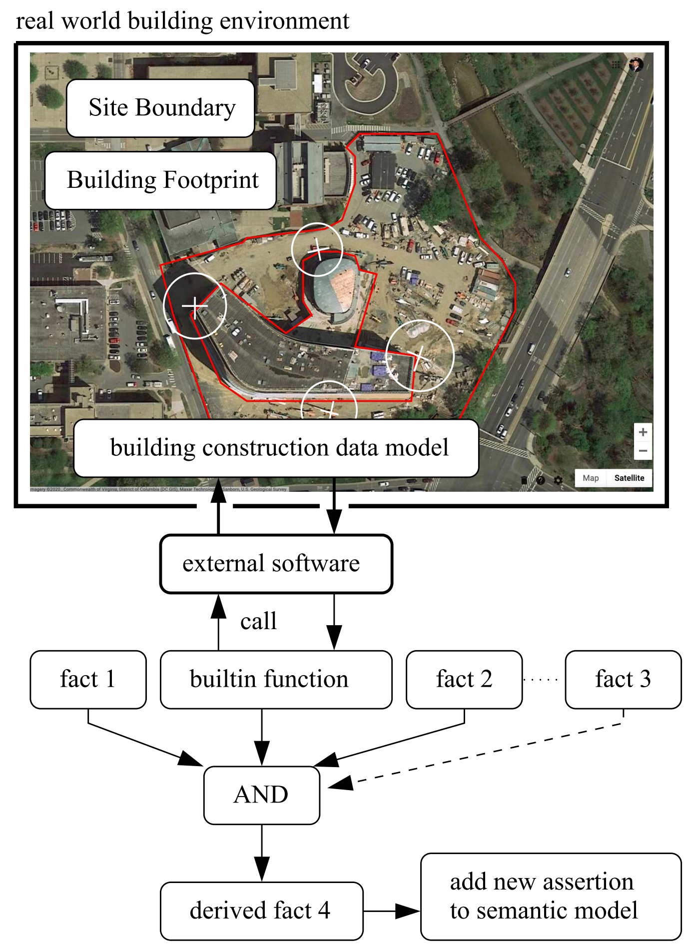


Working with Apache Jena and Jena Rules

Data-Driven Approach to Generation of Individuals in Semantic Graphs



Forward Chaining of Facts and Results of Built-in Functions to New Assertions…



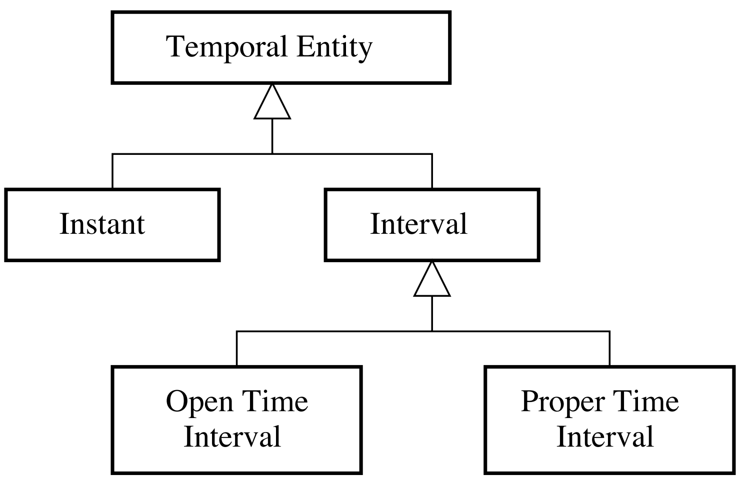
Meta-Domain Ontologies and Rules

Modeling Time

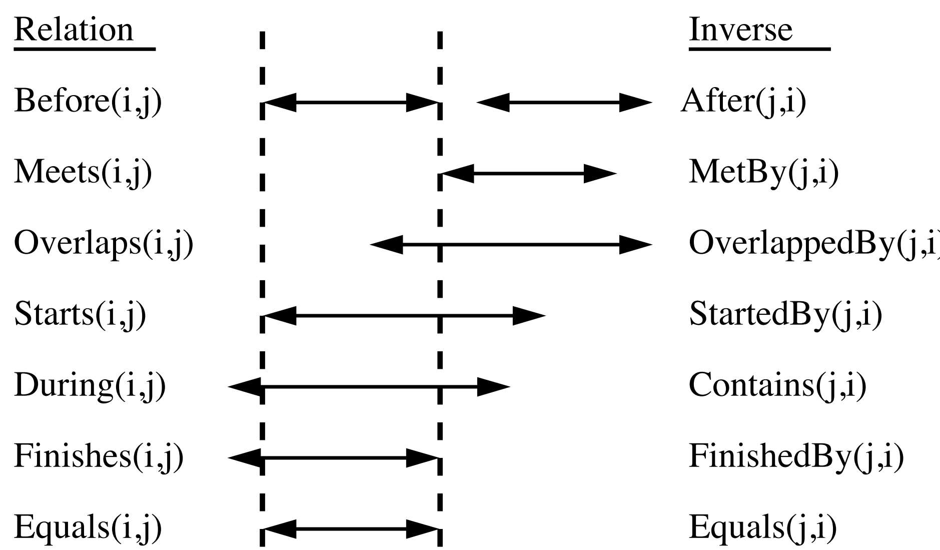
In order for a decision involving time to be reliable, the underlying models of time and theories of reasoning need to be formal.

Temporal Domain

(a) Ontology for Instants and Intervals of Time



(b) Logical Relationships Among Intervals of Time



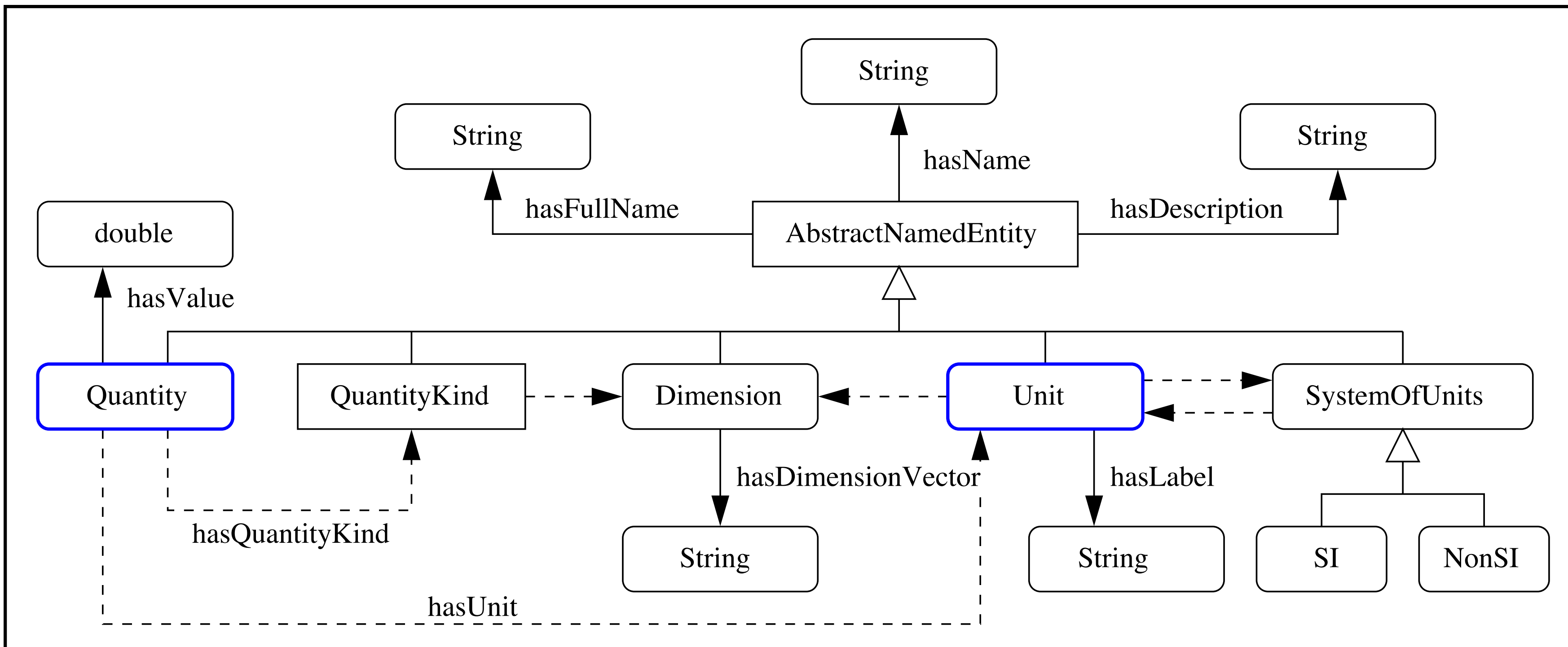
Source: Allen, 1983.

Meta-Domain Ontologies and Rules

Physical Units Domain: Length, Mass, time, Temperature

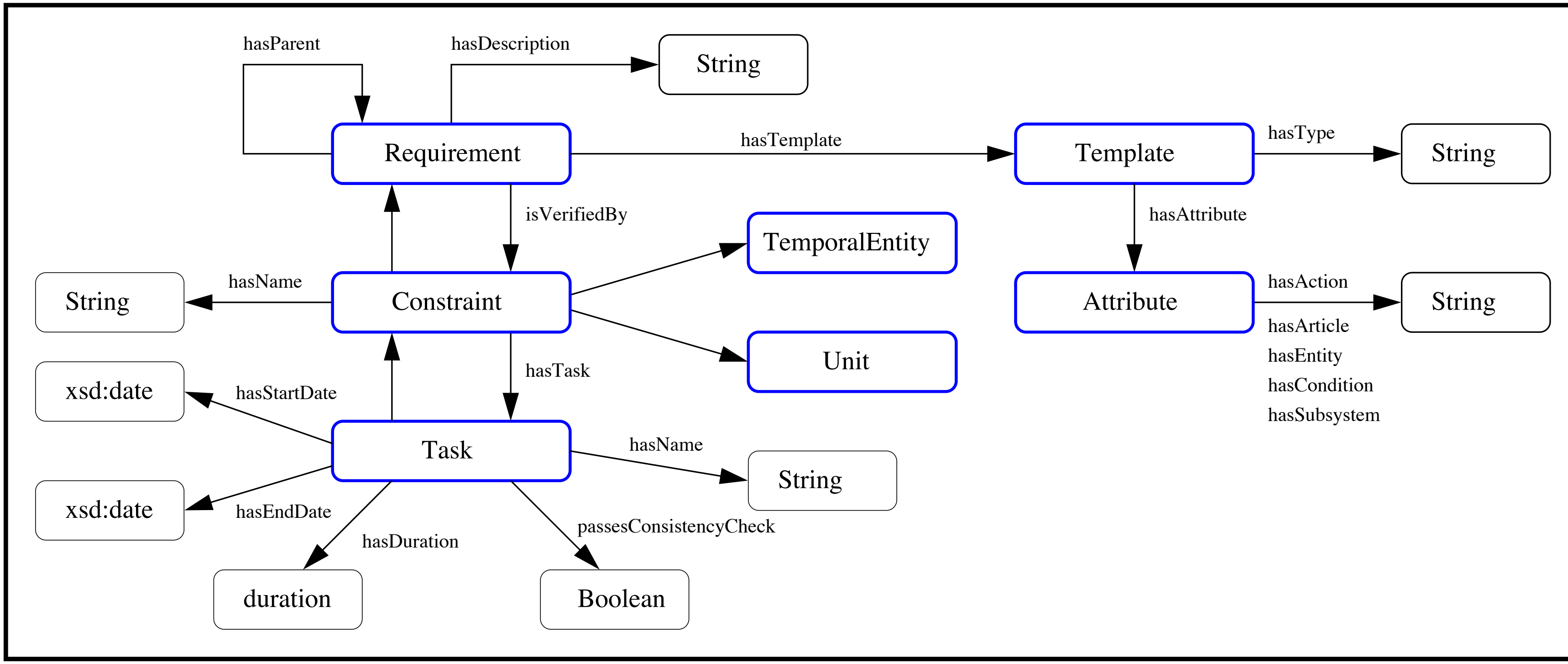
$$\text{unit} = k \cdot L^{\alpha} M^{\beta} t^{\gamma} T^{\delta} \cdot \text{rad}^{\epsilon}$$

Semantic Representation for Physical Quantities and Units



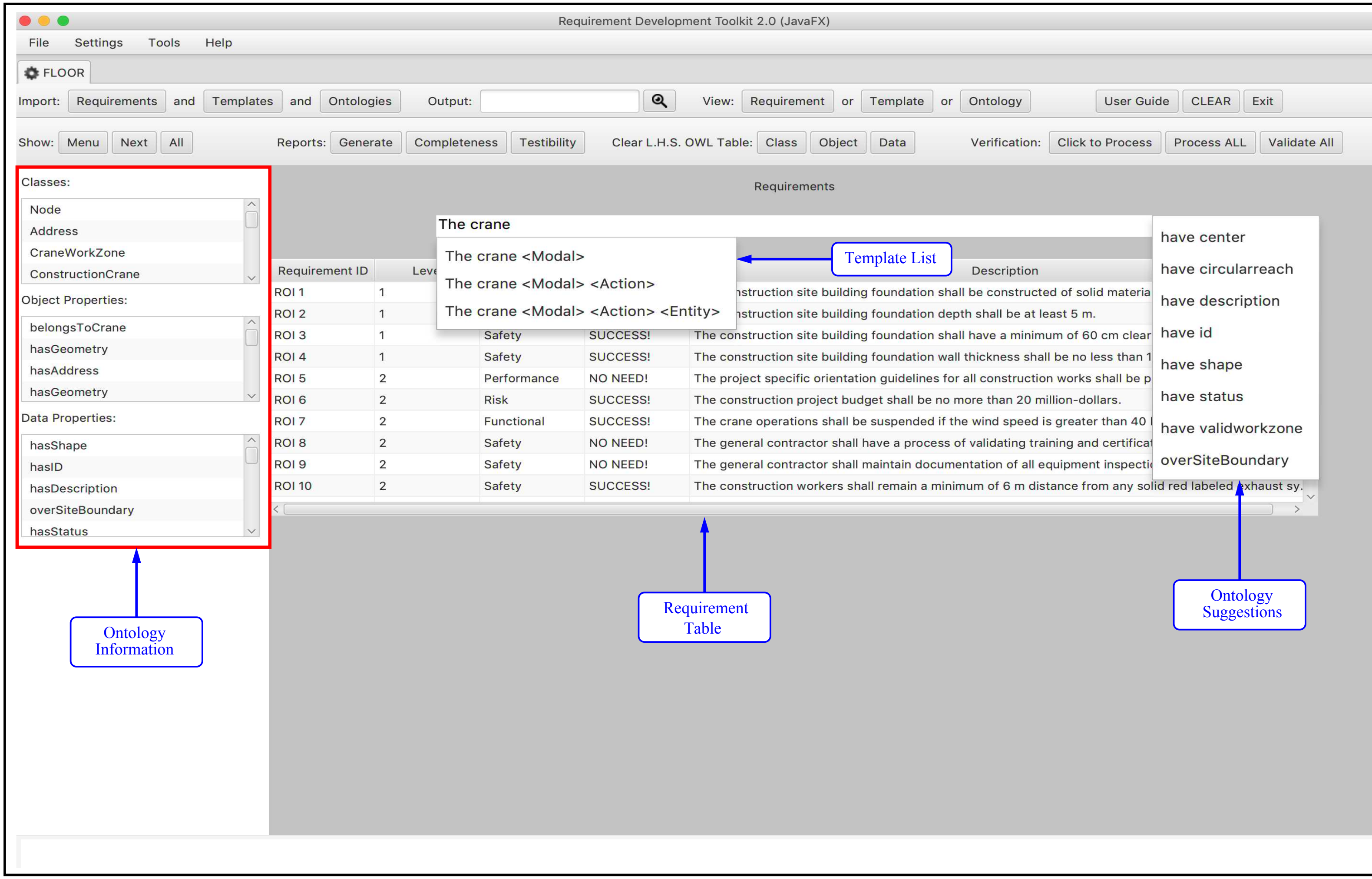
Domain-Specific Ontologies and Rules

Requirement, Constraint, Template, and Task Ontologies



FLOOR Software Prototype

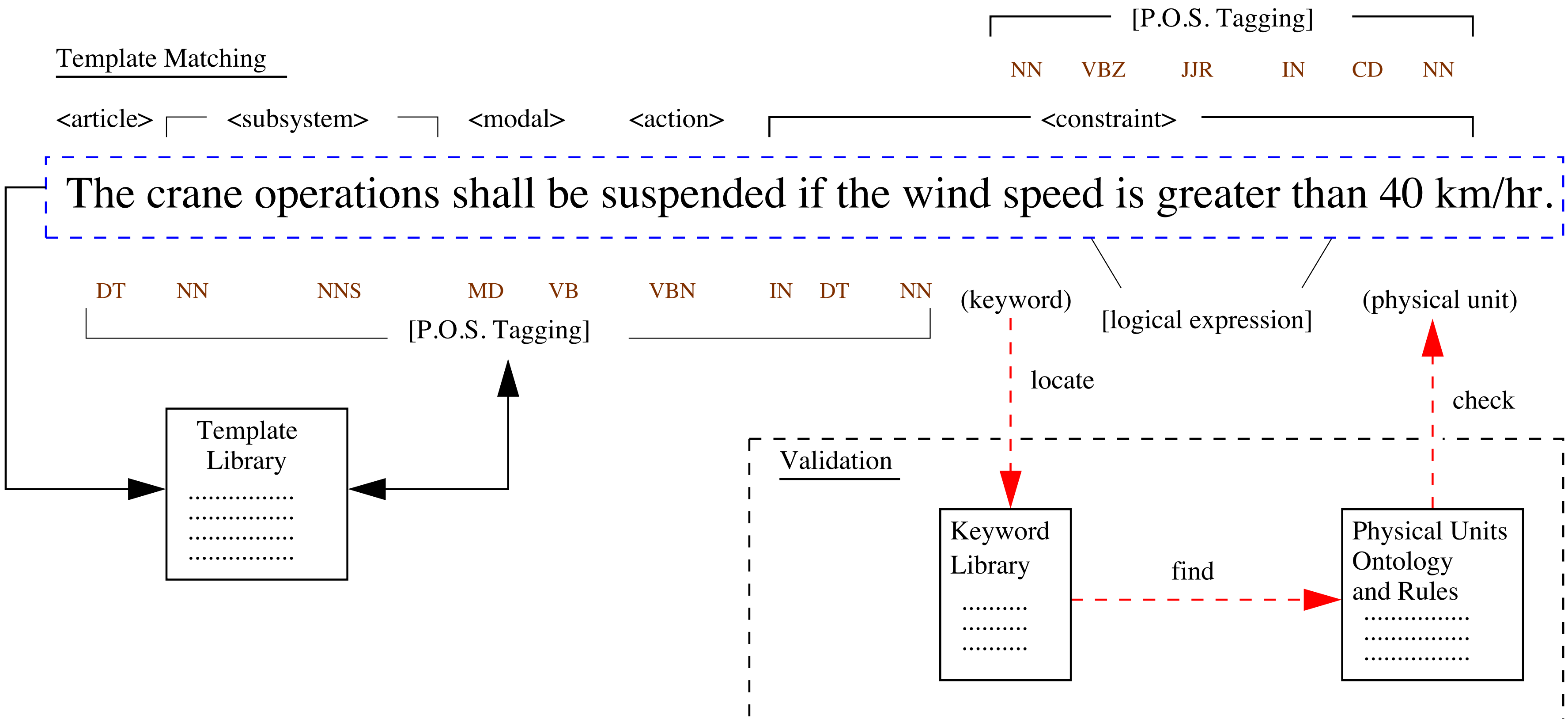
Framework for Linking Ontology Objects and Textual Requirements.



Semantic Processing and Validation

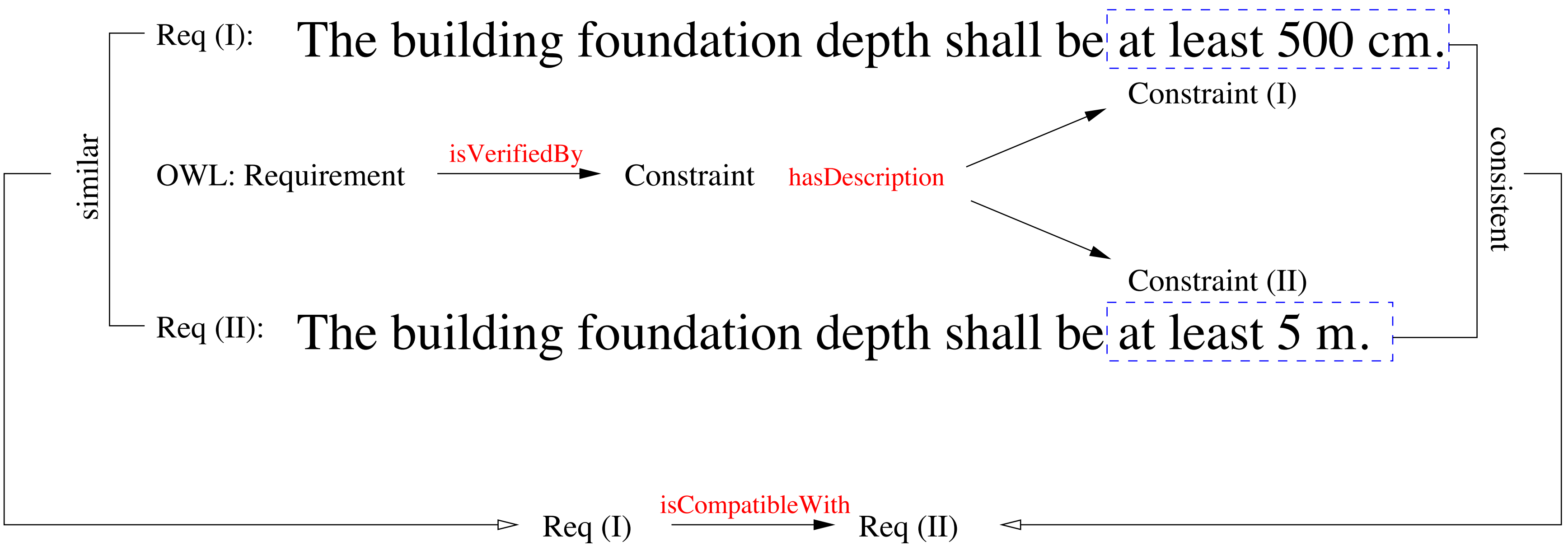
Case I: Detailed Validation of an Individual Requirement

Requirement: 5.2



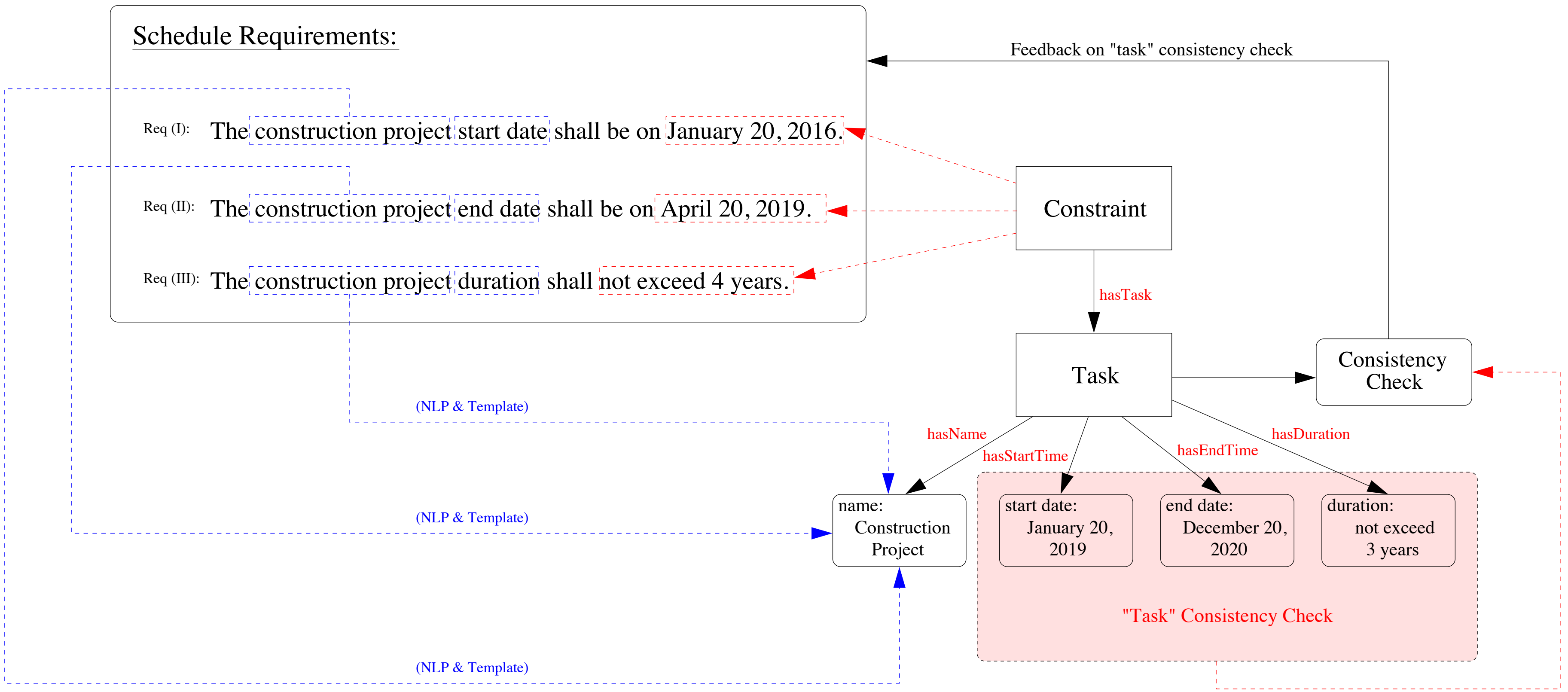
Case 2: Identifying Requirement Duplicates

Requirement: 3.2



Case 3: Group of Requirements Relating to a Single Task

Requirements: 2.1, 2.2 and 2.3



References

- **Allen J.F.**, Maintaining Knowledge about Temporal Intervals, Communications of ACM, Vol. 16, No. 11, 1983, pp. 832-843.
- **Apache Jena**, Open Source Java Framework for Building Semantic Web and Linked Data Applications, See: https://jena.apache.org. 2022.
- **Arellano A. Zontek-Carney E., and Austin M.A.**, Frameworks for Natural Language Processing of Textual Requirements, Int J. Adv. Systems and Measurements, Vol. 8, No 3., 2015, pp. 230-240.
- **Arellano A., Zontek-Carney E., and Austin M.A.**, Natural Language Processing of Textual Requirements, In Proc. 10th International Conference on Systems (ICONS 2015), 2015, pp. 93-97.
- **Borjigin S., Austin M.A., and Zontek-Carney E.**, Semiautomated Development of Textual Requirements: Combined NLP and Multidomain Semantic Approach, Journal of Management in Engineering, ASCE, September, 2022.
- **Coelho M., Austin M.A., and Blackburn M.R.,** The Data-Ontology-Rule Footing: A Building Block for Knowledge-Based Development and Event-Driven Execution of Multi-Domain Systems, 2018 Conference on Systems Engineering Research, Charlottsville, VA, May 8-9, 2018. Also see: Chapter 21, Systems Engineering in Context , Springer, 2019.
- **Coelho M., Austin M.A. and Blackburn M.R.,** ``Semantic Behavior Modeling and Event-Driven Reasoning for Urban System of Systems,'' International Journal on Advances in Intelligent Systems, Vol. 10, No 3 and 4, December 2017, pp. 365-382.
- **Zontek-Carney E.**, Framework for Linking Ontology Objects and Textual Requirements, MS Thesis in Systems Engineering, University of Maryland, College Park, May 2017.

Thank You

Questions?

Contact Information

Mark Austin: austin@umd.edu,
http://www.isr.umd.edu/~austin