SYSTEMS ENGINEERING RESEARCH CENTER | AIRC ACQUISITION INNOVATION RESEARCH CENTER | SERC RESEARCH REVIEW 2026

Natural Language Processing of Acquisition Text

WRT-2604.03 AI Workforce Support Tools to Speed DFARS Rulemaking

Sponsors: Ms. Kimberly Ziegler, DPCAP/DARS, Ms. Nickee Abbott, OPDASW(A)

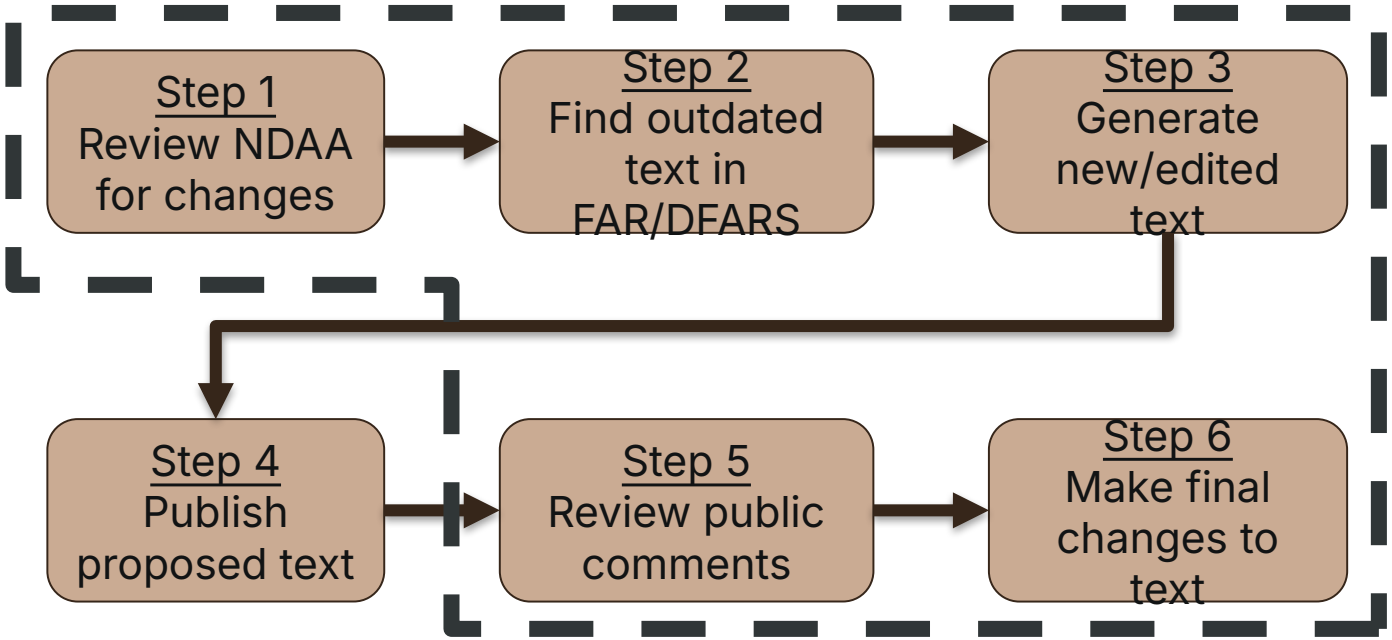
Dr. Jose Ramirez-Marquez

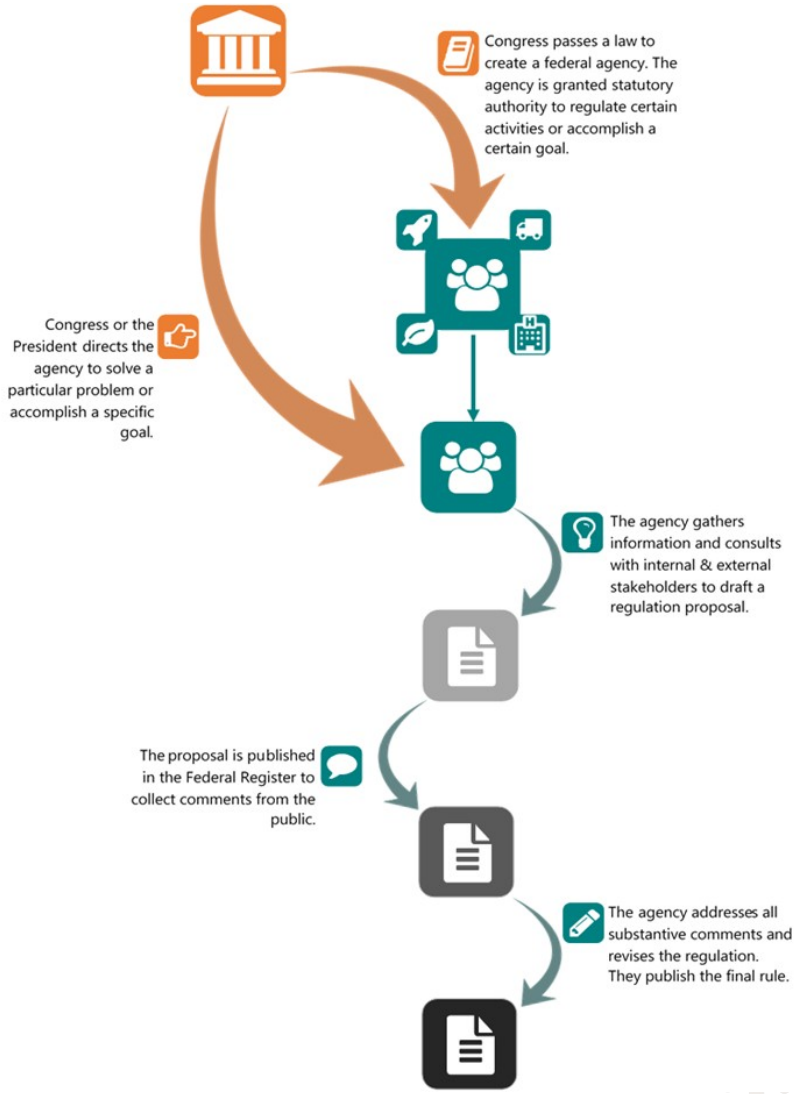
STEVENS INSTITUTE of TECHNOLOGY THE INNOVATION UNIVERSITY 1870

VIRGINIA TECH

Current State of Acquisition Regulation Rulemaking

The manual monitoring and analysis of text is highly time-consuming, hindering effective decision-making and rapid implementation of requirements.





Why NLP for Acquisition?

Acquisition decisions depend on extensive statutes, regulations, policies, and technical documents.

Challenge: Visual situational awareness of interactions and impacts.

Manual review makes it difficult to trace relationships, recognize emerging themes, and prioritize changes.

Challenge: Systematic review of evidence, quality, and outcomes.

Solutions: NLP and AI transform unstructured text into searchable, comparable, and interpretable evidence.

Related AIRC research work: WRT-1079 (J-Book keyword association); WRT-2507.11A (document analytics); WRT-2507.11B (NDAA–DFARS analysis); WRT-2604.03 (AI support for DFARS rulemaking).

A Unified Research Perspective

TEXT ANALYTICS — ACROSS DOCUMENTS

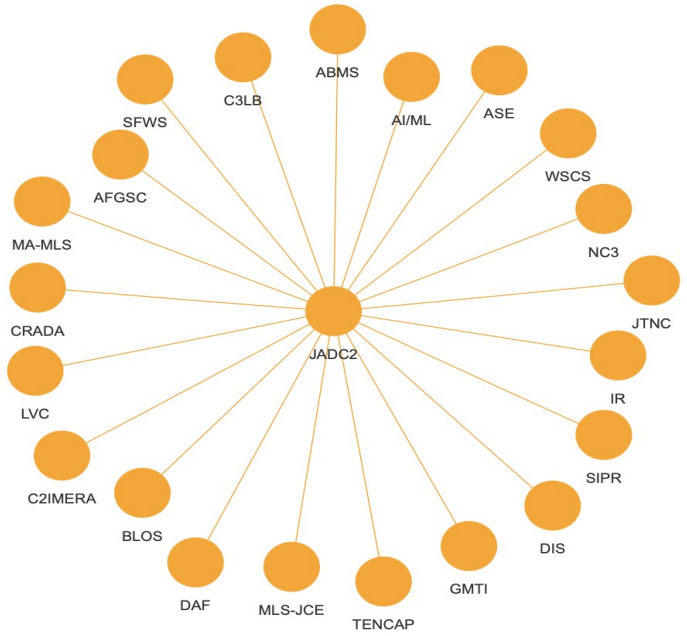
- **Extract relevant passages and concepts**
  - Example: J-Book keyword search for JADC2
  - Impact: Faster retrieval and traceable concept associations
- **Identify keywords, entities, and themes**
  - Example: NLP extraction and similarity analysis
  - Impact: Discover related concepts across documents
- **Summarize and classify documents**
  - Example: Acquisition document analytics portal
  - Impact: Faster synthesis and thematic review

DECISION ANALYTICS — ACROSS DOMAINS

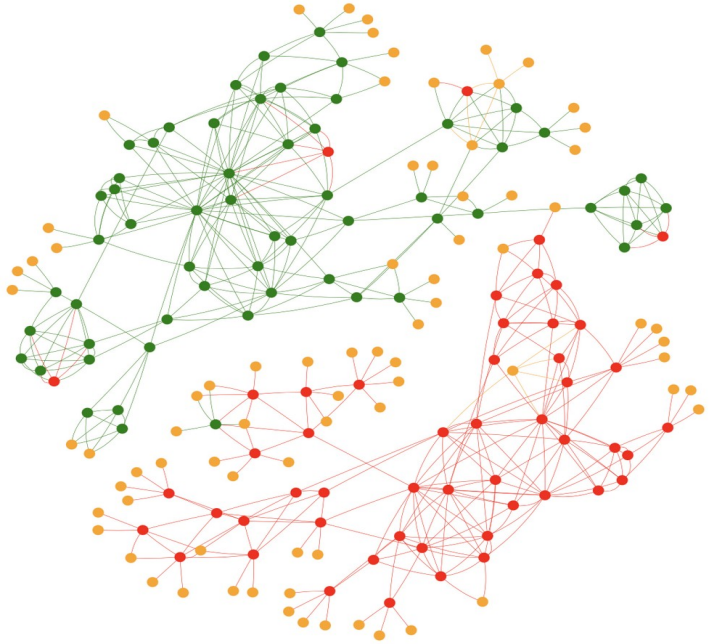
- **Connect related policy artifacts**
  - Example: NDAA sections linked to DFARS cases
  - Impact: Trace statutory-to-regulatory relationships
- **Compare regulations and propose changes**
  - Example: NDAA §817 and DFARS draft revisions
  - Impact: Analyst-supported change identification
- **Visualize patterns for human review**
  - Example: Document and skill relationship networks
  - Impact: Reveal communities and connections

Application 1 | J-Book Search and Association

From keyword extraction to interpretable concept networks



JADC2 and adjacent concepts



Iterative expansion of concept communities

Extracted keyword associations become network nodes and links, allowing exploration beyond an initial search term.

Application 2 | From NDAA Text to DFARS Updates

Identify relevant provisions, match regulatory text, and assess proposed edits

Example: NDAA 2022 §817 — repeal of preference for fixed-price contracts. The prototype compares generated DFARS edits with reference changes.

Table 1 Quality Index for Expected NDAA Sections Versus Extracted NDAA Sections

Fiscal Year	Extracted	Expected	Quality Index	Missed Section Numbers
2010	15	16	94%	745
2011	18	20	90%	834, 893
2012	26	28	93%	841, 834
2013	18	20	90%	802, 1651
2014	7	7	100%	
2015	15	16	94%	745
2016	25	26	96%	1214
2017	41	43	95%	873, 1212
2018	30	31	97%	1214
2019	30	31	97%	888
2020	31	31	100%	
2021	28	29	97%	832
2022	20	20	100%	
2023	19	19	100%	

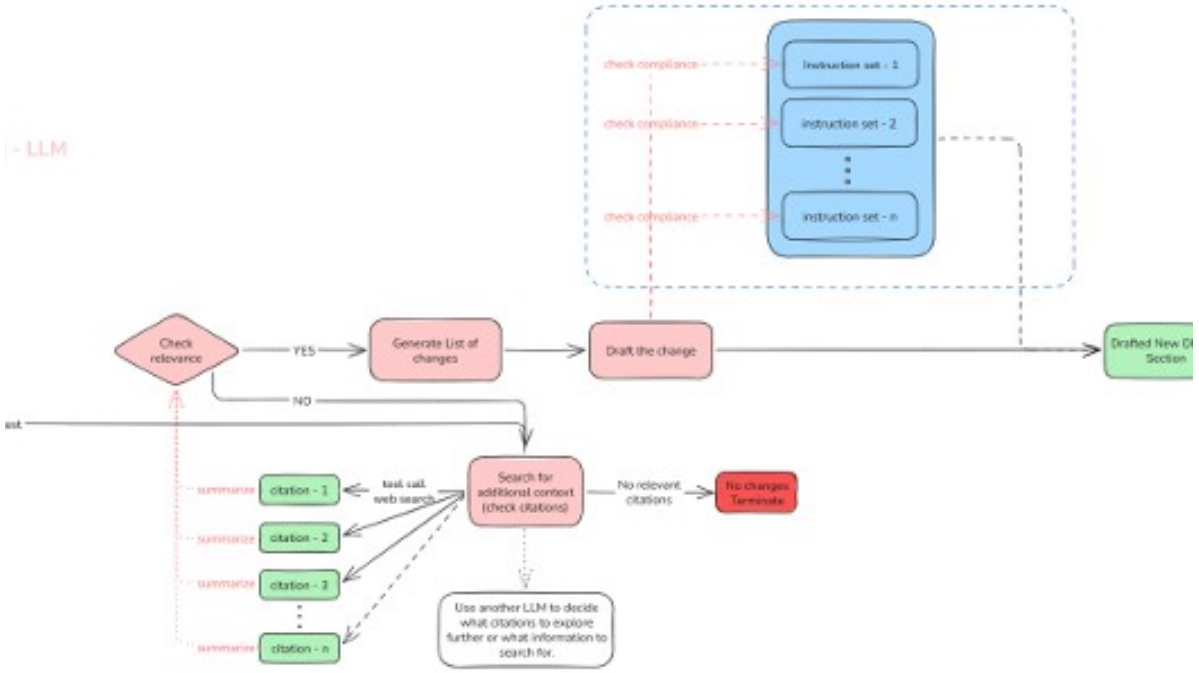
Table 2. Quality Index for NDAA Section to FAR/DFAR Section(s) Mappings

Fiscal Year	Sections w/ Published FAR/DFARS Mappings	Sections <50% Mapping Coverage	Quality Index (%)
2010	15	2	87%
2011	20	3	85%
2012	31	7	77%
2013	24	7	71%
2014	7	1	86%
2015	18	7	61%
2016	27	6	78%
2017	50	16	68%
2018	33	11	67%
2019	28	9	68%
2020	31	5	84%
2021	29	9	69%
2022	20	6	70%
2023	19	5	74%

Sources: Jose Ramirez-Marquez et al., 2026 (NPS ARS)

Generation of Draft Regulatory Text Changes

- Evaluate the need for text changers using an LLM (GPT 4.1) agent
- Collect additional context with a 2nd LLM
- Determine the necessary changes using a 3rd LLM given the input sections and the identified context
- A 4th LLM drafts modified regulation text using FAR drafting guidelines and the necessary changes



Process for Generating Changes to a DFARS Section Given an Old DFARS Section and an N Section

Prototype now in beta-testing and being transitioned to AIRC's DARCIE CUI-capable analytic environment

Application 3 | End-to-End Document Intelligence

Document ingestion → extraction → NLP analysis → relationship discovery → analyst review

1 UPLOAD
PDF / text corpus

2 PARSE
Extract text

3 ANALYZE
Keywords, similarity, AI summaries

4 DISCOVER
Clusters and networks

5 REVIEW
Human interpretation



Analytical outputs

- Document similarity and keyword insights
- AI-generated summaries
- Thematic clusters and network visualization
- Source-linked analyst review

Demonstration: 244 executive orders processed end-to-end; approximately 19 minutes reported.

Related Application | Organizational Skill Networks

Collect → Extract and normalize → Build skill networks → Benchmark → Identify gaps

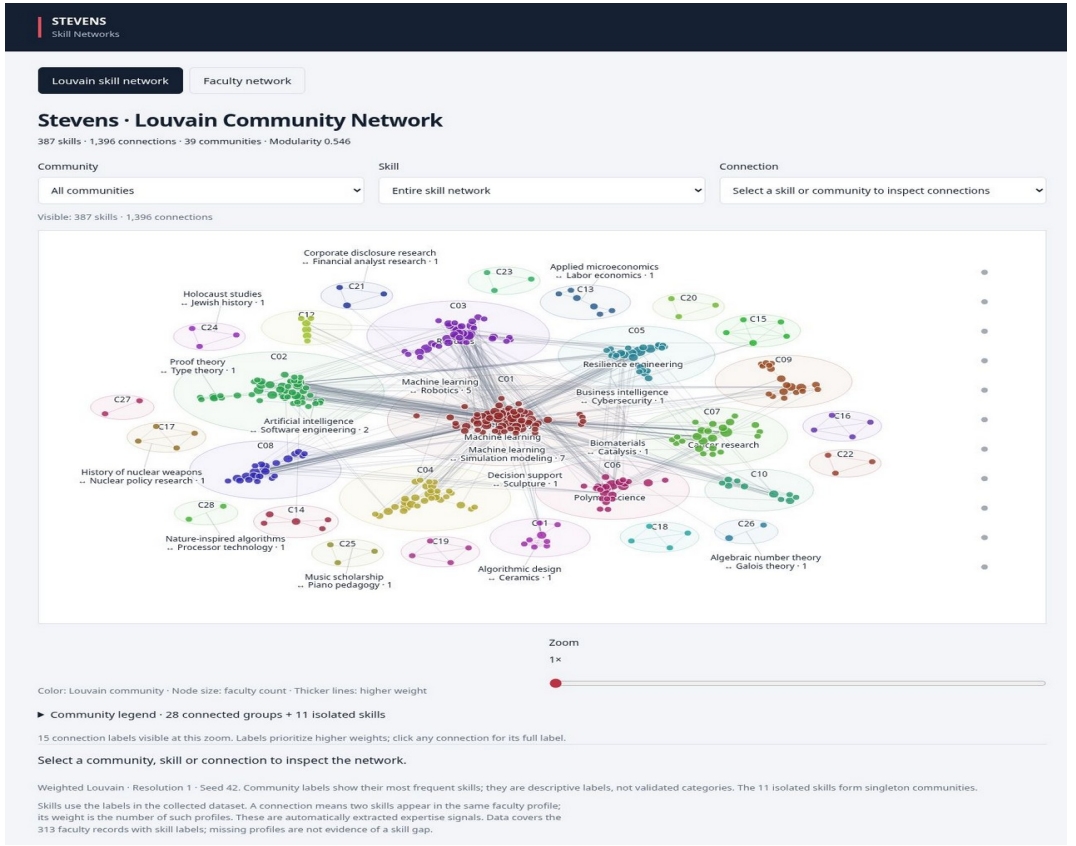
Stevens faculty pilot

313 faculty with skill labels

387 skills and 1,396 weighted connections

28 connected communities; 11 isolated skills

Evidence of skill relationships, not proof of individual proficiency.



Interactive pilot: stevens-faculty-skill-network.linktera-rob-9535.chatgpt.site

Research Challenges, Next Steps and Takeaways

Research challenges

- Heterogeneous source formats
- Semantic ambiguity and evolving terms
- Traceability and reproducibility
- Hallucination and expert validation

Future directions

- Cross-document policy intelligence
- Impact analysis for regulatory changes
- Integrated human-in-the-loop review
- Scalable dashboards and workflows

KEY TAKEAWAY

NLP connects acquisition evidence across documents. Combining extraction, reasoning, and networks can support decision-making, while expert validation remains essential.

SYSTEMS ENGINEERING RESEARCH CENTER | SERC RESEARCH REVIEW 2026

Thank you!

Email the presenter

✉ Jose Ramirez Marquez, jmarquez@stevens.edu

Other research team leads

Brian Mayer	Virginia Tech
Naren Ramakrishnan	Virginia Tech
Doug Buettner	Stevens

Stay connected with SERC online