STEVENS INSTITUTE of TECHNOLOGY
THE INNOVATION UNIVERSITY®

Analyzing and Assessing Contracts for Embedded Risk

Sponsor: DAU

Dr. Carlo Lipizzi

13th Annual SERC Sponsor Research Review
HOSTED VIRTUALLY ON: November 3, 2021
www.sercuarc.org

WRT-1023

- **Title**: Analyzing and Assessing Contracts for Embedded Risk
- The goal of this research effort is to apply data analytics to understand the assessment processes undertaken by a contracting officer. The intent is to bring significant efficiencies to these assessment processes and develop a prototype tool covering relevant parts of the DoD contracting process from beginning to end

WRT-1023

The effort has been divided into several phases. The researchers is focused on:

- Determining the most appropriate contract structure, i.e., contract type, for the government to use in response to a specific set of requirements (**Phase 1**)
- Assessing the description of requirements, contract text, and/or a general question regarding a government solicitation or contract against the Federal Acquisition Regulations (FAR), Defense Federal Acquisition Regulations Supplement (DFARS), DFARS Procedures, Guidance and Implementation (PGI), DoD Class Deviations, and other applicable DoD guidance. (DoD Source Selection Procedures; DoD Risk Management Guide; etc., (www.dau.mil/tools) (**Phase 2**)
- Performing an after-triage contracts comparative analysis, limited to the compliance to formal and critical issues (**Phase 3**)

WRT-1023: Key features

- Leveraging on current MO and literature, create a logical framework to classify requests based on given Contract Types
- Create a computational model for the logical framework
- Create a visualization systems to present the results
- Deliver the results with an agile approach, developing prototypes/proofs of concepts with increasing capabilities
- As per our contract, during the first year we developed an early prototype to prove validity of the approach. The prototype covers the basic functionalities highlighted above but with limited robustness, interactivity, proactivity and reusability

The context: scenario

- We have been informed of the existence of 10 different contract types, some of them with relevant degrees of similarity
- The job of the contracting officer we are addressing with this research task is to classify acquisition requests by contract type
- Working with the Sponsor, we identify keywords that are characteristic of each contract type. No plain combination of those keywords leads to an exact classification
- The classification is rarely a black and white decision and is mostly based on the knowledge and experience of the contracting officer

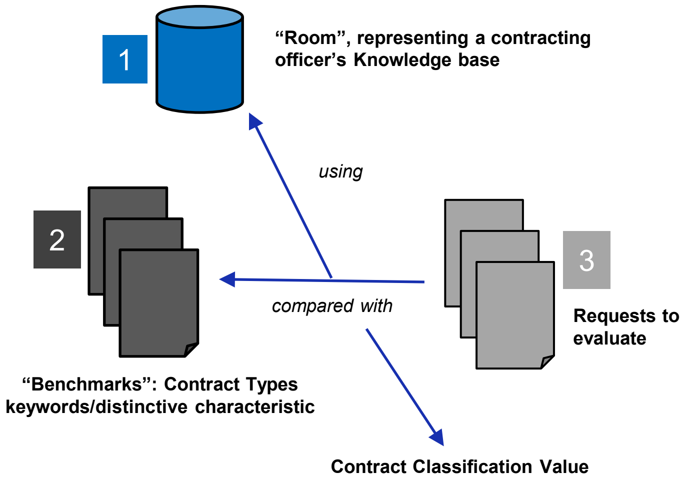
What shaped the system

- The classification we want to implement is rooted in a semantic analysis of text
- The semantic value in this case is the proximity of a request to a contract type, according to contracting officers' knowledge
- It was essential to “recreate” the contracting officers' knowledge to be used for the classification
- We leveraged on what we did in WRT-1010 to create a computational representation of that knowledge base

WRT-1023 key component: the “room theory”

- The ”room theory” is a framework to address the relativity of the point of view by providing a computational representation of the context we want to use to evaluate the text
- The non computational theory was first released as “schema theory” by Sir Frederic Bartlett (1886–1969) and revised for AI applications as “framework theory” by Marvin Minsky (mid ‘70)
- For instance, when we enter a physical room, we instantly know if it is a bedroom, a bathroom, or a living room
- Rooms/schemata/frameworks …
    - Are mental frameworks that an individual possesses
    - A mental framework is what humans use to organize remembered information
    - Represent an individuals view of reality and are representative of prior knowledge and experiences
- We create computational ”rooms” by processing large corpora from the specific domain/community generating numerical dataset (“embeddings table”). We consider a table as a knowledge base for the context/point of view
- The “room” method makes the whole approach easy to be moved to different domains

How the “room theory” works



- “**Room theory**” enables the use of context-subjectivity in the analysis of the incoming documents
- Context-subjectivity can be the point of of view of a subject matter expert
- The context-subjectivity in the analysis is represented by a domain specific numerical knowledge base, created from a large domain specific & representative corpus that is then transformed into a numerical dataset (“embeddings table”)

- The key components are:

1. A point of view for the comparison (the “room”). This is represented by the embeddings table extracted from a large/representative corpus from the specific domain
2. A criteria for the analysis (the ”benchmark”). This is a list of keywords defining the “what we are looking for”. Different benchmarks would provide different analyses

Implementation challenges

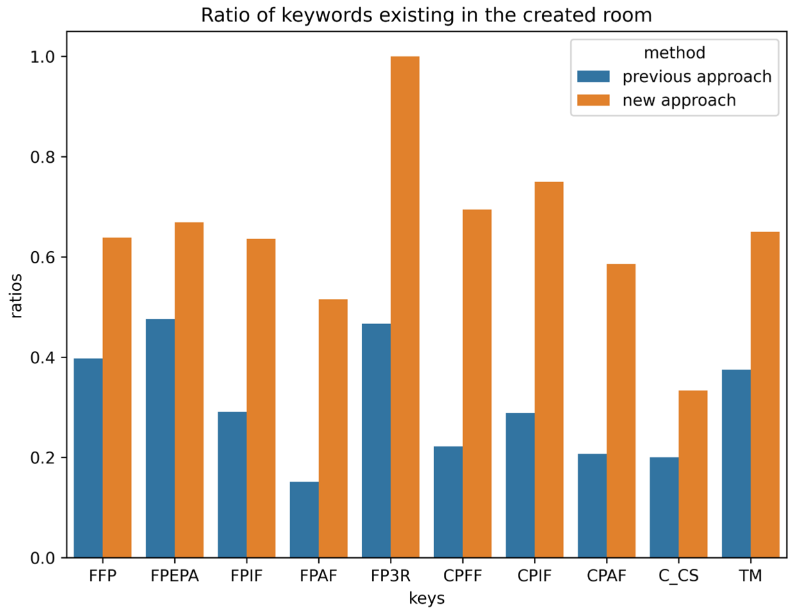
- The initial ~100 keywords were not enough to define the contract types
- The initial ~200 documents/~30K unique words were not enough to create a knowledge base able to "understand" the requests. The knowledge base was based on a corpus created from contracting-related text
- The "words" are rarely just "words", being most of the times combination of words, like "ceiling price" or "time and materials"
- Not all the words have the same relevance in determining a contract type
- The "similarity" between request and contract type is calculated by evaluating distance between vectors. Standard methods (like cosine similarity) average the results
- The user cannot work with Python code to have their requests processed

Implementation answers

- We now have 464 keywords. This includes the synonyms of the original keywords
- We expanded the sources including related domains. The corpus representing a contracting officer’s knowledge base is now composed by 537 documents, for a total of 119,941 unique words
- We developed a domain-agnostic method to extract the “semantic chunks” (“n-grams”) from the corpus. With this method we created the domain-specific “chunks”
- With the great contribution of our Sponsor, we assigned weights to each keyword
- We used “word mover” approach to calculate the vector distances
- We created a user interface to let the user input their documents/requests to be processed

Implementation answers: the impact of the upgrades

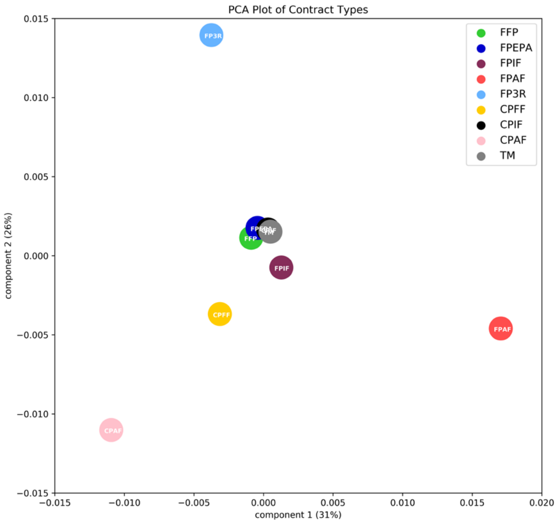
Coverage of the room

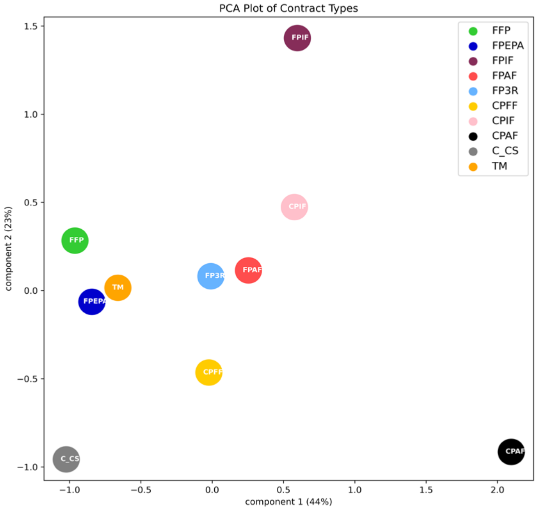


- The ~200 documents corpus generated a knowledge base not able to identify the keywords either properly or at all
- The new one based on 537 documents is performing much better

Implementation answers: the impact of the upgrades

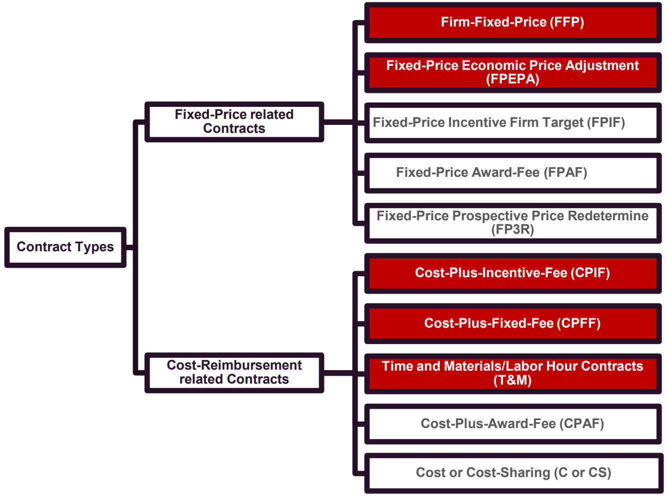
Differentiation capability





- With the new keywords and weights, the system is able to better differentiate between contract types

Classification results



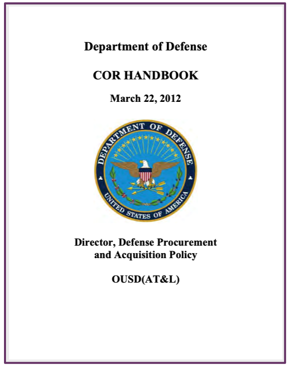
	doc_name	FFP	FPEPA	FPIF	FPAF	FP3R	CPFF	CPIF	CPAF	C_CS	TM	Predicted Contract Type
0	W90VN920T0002 Security Containers.pdf	43.17%	10.11%	0.18%	4.97%	2.88%	3.03%	0.25%	3.73%	4.45%	27.24%	FFP
1	Attch 11-Past Performance Questionnaire-PPQ.pdf	1.70%	14.21%	0.12%	4.50%	55.46%	1.75%	0.67%	0.97%	20.25%	0.37%	FP3R
2	SOW.pdf	25.53%	29.06%	0.74%	2.34%	9.70%	4.21%	1.62%	0.19%	25.20%	1.42%	FPEPA
3	Sources Sought Notice_Revised.pdf	4.72%	29.93%	0.21%	0.36%	8.54%	2.95%	0.97%	0.09%	51.90%	0.32%	C_CS
4	SPRHA420Q0267.pdf	75.93%	12.27%	0.06%	0.07%	0.79%	0.96%	0.37%	0.21%	2.07%	7.27%	FFP
5	Solicitation Amendment FA527020Q00050001 SF 30...	6.08%	40.38%	0.10%	1.75%	8.15%	0.89%	0.62%	0.62%	41.12%	0.29%	C_CS
6	Solicitation# W9124720T9000.pdf	60.84%	13.48%	0.82%	0.50%	2.35%	0.20%	2.08%	0.10%	3.07%	16.56%	FFP
7	Questions and Answers for FA527020Q0005.pdf	4.94%	40.31%	0.81%	1.10%	9.13%	1.09%	1.53%	0.14%	38.72%	2.22%	FPEPA
8	Attch 13-Consent Form.pdf	5.01%	8.38%	0.13%	3.55%	23.07%	2.59%	0.70%	2.23%	54.14%	0.19%	C_CS
9	COMBINED SYNOPSIS SOLICITATION N3220520Q0159.pdf	71.99%	11.86%	0.06%	0.71%	7.67%	0.17%	0.31%	0.10%	5.73%	1.41%	FFP
10	Exhibit 4 - DHA Form 49.pdf	4.80%	21.28%	0.14%	2.84%	8.68%	1.94%	1.17%	1.96%	56.84%	0.36%	C_CS
11	SP7000-20-Q-1012.pdf	77.14%	11.09%	0.06%	0.28%	1.76%	0.55%	0.21%	0.21%	1.53%	7.17%	FFP
12	Domitory of Work Attachment 1 - Statement of W...	11.10%	25.56%	0.87%	3.08%	12.82%	2.34%	2.13%	0.22%	34.86%	7.03%	C_CS

- The ability of the system to differentiate between contract types is calculated using PCA and performing T-tests for “within group distances” vs. “between groups distances” and then running Monte-Carlo Simulations to generate a validity report

Use Case: Individual Documents

Each individual artifact is treated as a separate document

- The user wants to upload several different Requests for Proposals to see which category of contract type should be strongly considered for each separate document
  - Streamline work
  - Allows potential, or new, Contracting Officers to familiarize themselves with the selection and analyzation process
  - There can be several Requests for Proposals for one specific job. A DoD employee may want to see how each document is worded to attract the specific Cost Proposals
  - Evaluate the risk associated with each artifact by understanding the contract types
  - A DoD employee may also want to comprehend the roles and responsibilities for contract surveillance
- Potentially, the Contracting Officer and/or Government Employee can:
  - Have specific risks highlighted for the various documents.
  - Provide feedback so the prototype can be evolved using reinforced learning



Department of Defense

COR HANDBOOK

March 22, 2012

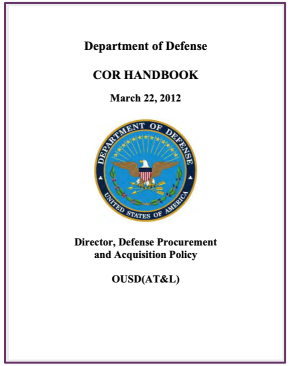
Director, Defense Procurement
and Acquisition Policy

OUSD(AT&L)

Use Case: Concatenated Documents

Multiple artifacts uploaded are treated as one document

- The user wants to upload several different artifacts (Requests for Proposals, etc.) to see which category of contract type should be strongly considered if treated as one document.
  - Streamline work
  - Allows potential, or new, Contracting Officers to familiarize themselves with the selection and analyzation process for an overall objective
  - Again, there can be several Requests for Proposals for one specific job. A DoD employee may want to see various documents (including ancillary) linked together to determine
  - A DoD employee can use the tool to help keep files current and complete, understanding the contract type that the documents support
- Potentially, the Contracting Officer and/or Government Employee can:
  - Provide feedback so the prototype can be evolved using reinforced learning
  - Move from a prototype to a Minimum Viable Product (MVP)
  - Eventually highlight potential risks factors within the texts as well as the various aspects that drove the entirety of the documents to be deemed the specific contract type. It also can aid in managing potential problem areas



Department of Defense

COR HANDBOOK

March 22, 2012

Director, Defense Procurement
and Acquisition Policy

OUSD(AT&L)

Demo

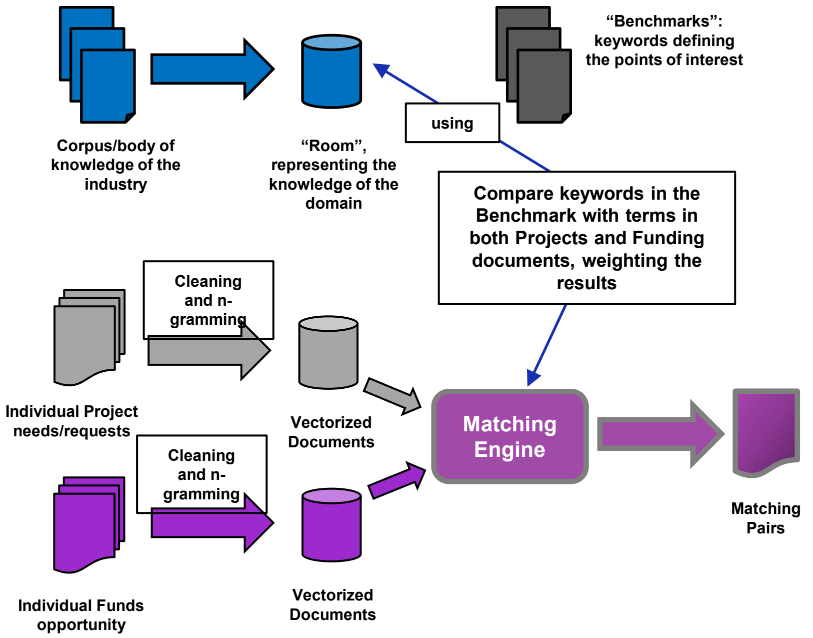
Future Developments

- The current “system” is a proof of concept, with known limitations, not tested on larger scale, not incorporating feedback from users
- Benchmark the results on a larger scale, using existing cases
- Introducing an after-triage supervised system to refine the results. This may be as refined/complex as needed, along 3 main lines:
  - Force case-specific behaviors
  - Add a layer of reinforcement learning, using the users feedback (accuracy of the results) as score to maximize
  - Take into account external factors, such as critical relevance of the request, previous agreements with preferred providers, competitiveness of the market segments (e.g.: fewer competitors, more risk)
- Expand the User Interface providing more interactive functionalities and visualizations
- Using the system as a learning tool, to provide younger contracting officers a way to evaluate/compare their choices

Derived use of the developments

- The overall approach of a system driven by a combination of a generic knowledge base of the domain, along with topic-specific elements of interest is applicable to a multitude of different cases
- Systems focused on Natural Language processing are becoming more and more popular, but the focus is more on “processing” than “understanding” and take actions accordingly
- Understanding Language is not “just” processing. Understanding is a human characteristic, analyzed by philosophers as part of Epistemology
- An accurate (by human standard) “understanding” can come only from a model of human mind
- The current leading models in NLP/”NLU” are focused on the algorithmic part, missing a real model representing how the knowledge is created and used. It is basically representing the brain, not the mind. The leading model for NLP (GPT-3 by Open-AI) has 175 billion parameters, feeding a neural network providing results as a black box

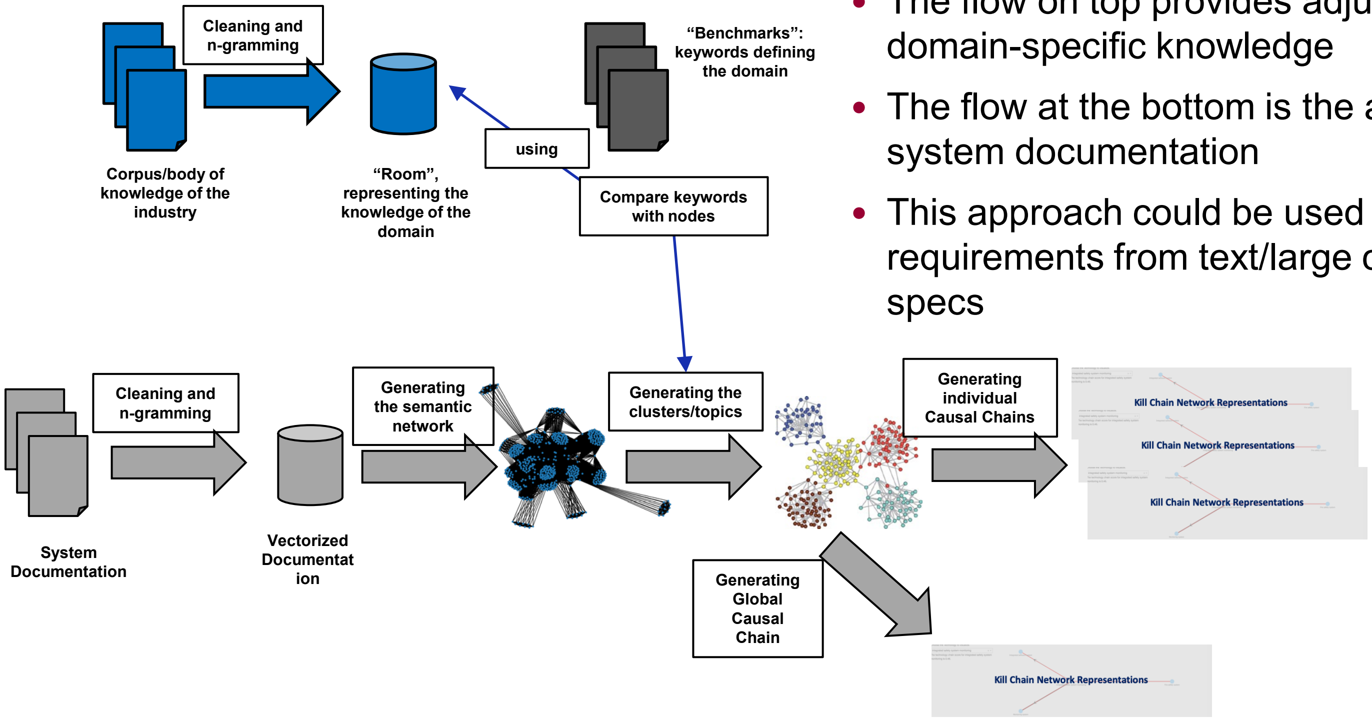
Example of application: matching projects and funds availability



This is a matching system driven by the domain knowledge and by specific weighted points of interest

- The "Room" is a computational representation of the knowledge in the given sector
- The "Room" is used to drive the matching based on knowledge of the sector
- The "Benchmark" is a list of keywords defining the points of attention for the match. Keywords will have weights to state their relevance
- This approach will address the need to have a match that is
    - Specific to the sector
    - Weighted by the predefined points of attention

Example of application: Extracting causal chains from text



- The flow on top provides adjustments based on domain-specific knowledge
- The flow at the bottom is the actual workflow on the system documentation
- This approach could be used to "reverse engineer" requirements from text/large corpora with systems specs

STEVENS INSTITUTE of TECHNOLOGY

1870

THE INNOVATION UNIVERSITY®

Thank you!

Dr. Carlo Lipizzi
clipizzi@stevens.edu

CENTER FOR COMPLEX SYSTEMS & ENTERPRISES