



A BOLD NEW FUTURE: THE PROMISE OF A DEFENSE ACQUISITION AI REVOLUTION

SAMANTHA NANGIA, ASN(RDA) DASN(P); THOMAS WARDWELL, ASN(RDA)
DASN(P); RANDALL MORA, AVUM INC

WHO WE ARE



Samantha Nangia

ASN(RDA) DASN(P)
Senior Program Analyst/
Procurement Technology
Transformation Lead



Thomas Wardwell

ASN(RDA) DASN(P)
Contracting Officer/ Financial Manager/
Acquisition Functional Area Manager



Randall Mora

Avum, Inc
CEO/ Lead Researcher/
Solution Architect

THE LIFE OF AN 1102











Could you help me write contracts in compliance with law and regulation?

6:32 AM ✓✓

I can provide guidance on contract structure and content, as well as help with researching relevant laws and regulations. However, I'm not a certified professional, and contracts for complex acquisitions like naval assets require expertise. Consider consulting a qualified contracting officer or attorney for specific guidance. Would you like some general information on contract writing or resources for further learning?

6:32 AM



What might be the best way you could help me if I had Acquisition and Contract documentation to review in my position at the deputy assistant secretary of the navy for procurement

6:34 AM ✓✓

Given your role, I can assist with:

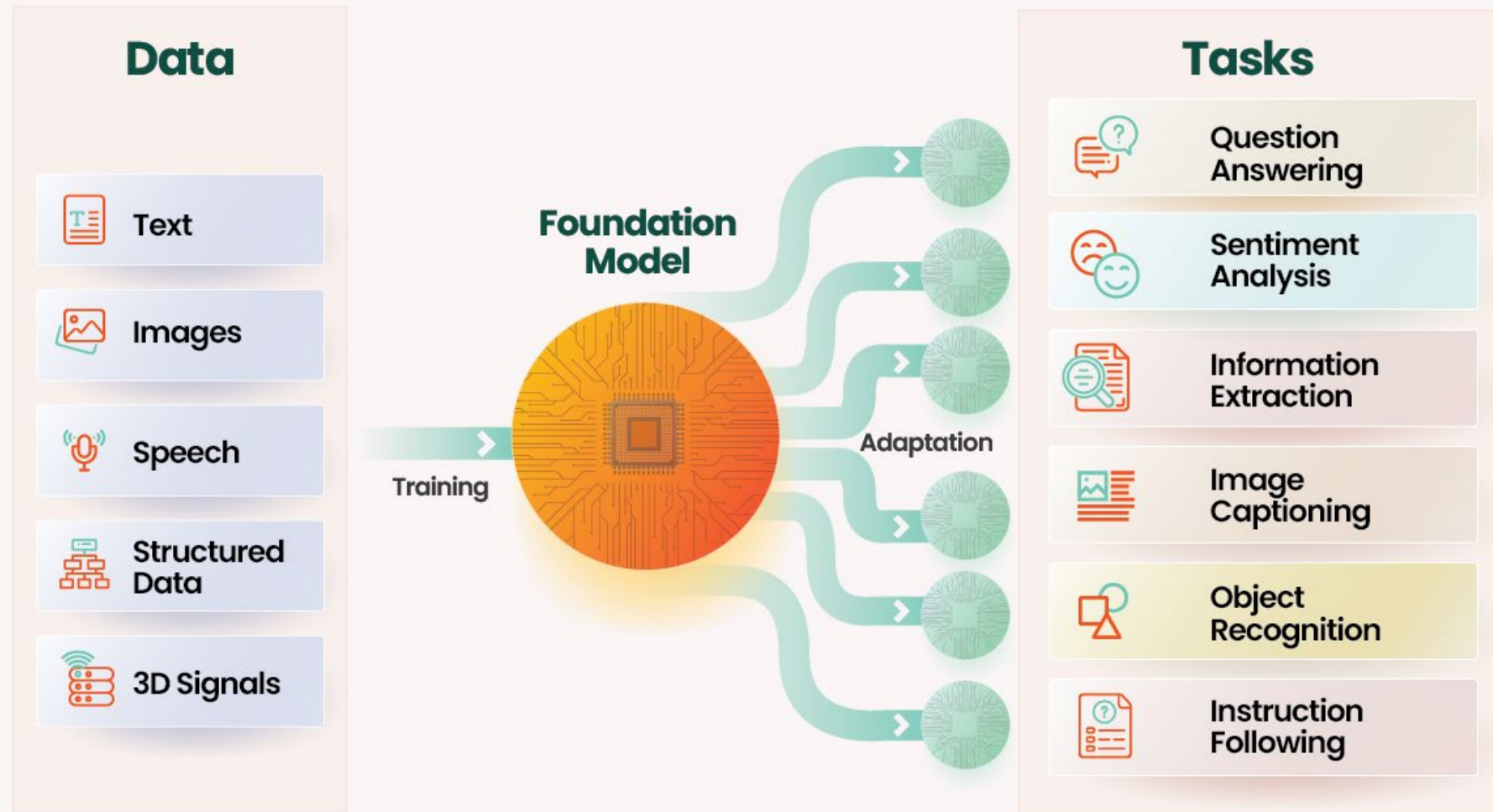
- Researching specific regulations, laws, or policies
- Providing guidance on contract formatting and structure
- Helping to identify potential issues or areas for review
- Offering suggestions for clarification or improvement

However, for actual review and approval of Acquisition and Contract documentation, it's essential to rely on qualified professionals with expertise in government contracting and procurement regulations.

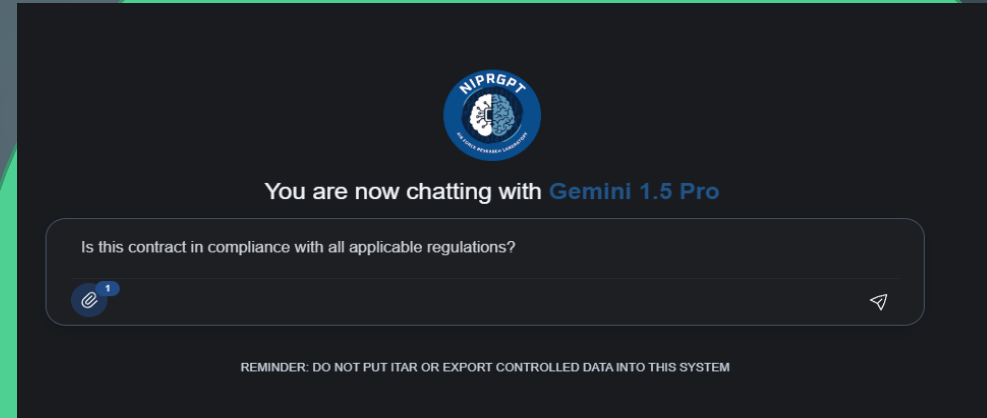
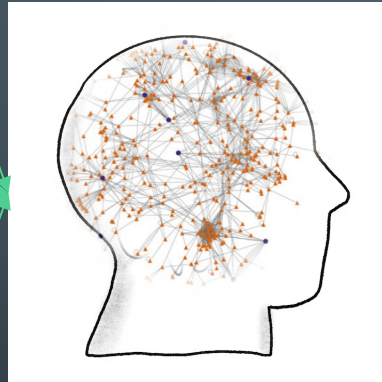
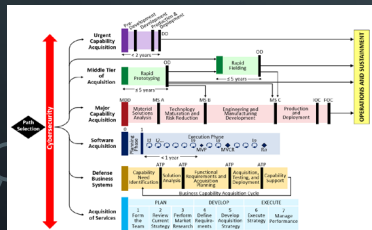
Would you like help with researching a specific regulation or policy, or perhaps guidance on a particular contract clause?

6:34 AM

How does Generative AI work?



DEFENSE ACQUISITION USE CASE



Contract Regulatory Compliance

UNCLASSIFIED

Key Takeaways for Compliance:

- **Cite Authority:** Clearly identify the authority under which the modifications are being made (e.g., Changes clause, bilateral agreement, etc.).
- **Detailed Justification:** Explain the rationale behind every change. Provide supporting documentation as needed.
- **Consistent Data:** Ensure all dates, values, and document references are accurate and consistent.
- **Clear Data Rights:** Address data rights explicitly and accurately.
- **Address Subcontracting:** Explain any changes to subcontracting plans and their compliance with small business goals.



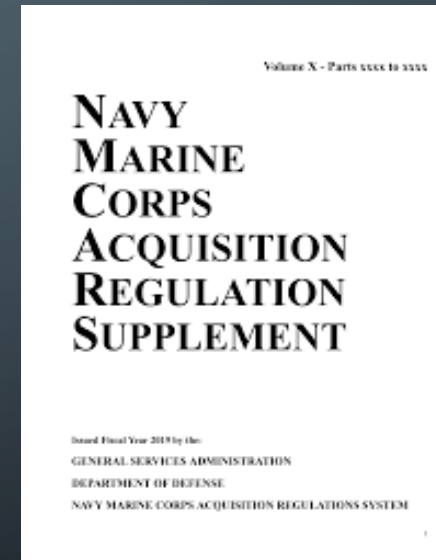
OUR SPECIFIC USE CASE

J&As and ISTRAPS come to DASN(P) for review as part of their routing to DASN(P)/ ASN(RDA) for signature

- W TAB A - NMCARS 18-25 Annex 1 - J&A
- W TAB B1 - CJA No. CR-24219 - Body - MS Word
- W TAB B2 - CJA No. CR-24219 - Track Changes - DASN(P) and AGC
- W TAB B3 - CJA No. CR-24219_Final Clean 22JAN25
- W TAB C1 - Draper CPS JA23-51 Final CLEAN
- W TAB C2 - 20240806_ASN(RD&A) CR_TC_Draper CPS JA23-51 Final CLEAN
- W TAB C3 - Draper CPS JA23-51 Final CLEAN 9.10.2024
- W TAB D1 - NAVSEA JA_DD91_DMP2_to_NASSCO
- W TAB D2 - ASN(RDA) CR_TC_NAVSEA_DD91 DMP J&A - NASSCO - includes Legal Review
- W TAB D3 - Final USS PINCKNEY (DDG 91) DMP JA 42,916_20231215

- W TAB A - NMCARS 18-25 Annex 18 - ISTRAP
- W TAB B1 - ISTRAP CNO Avails Fast Attack Subs
- W TAB B2 - ASN(RDA) CR_TC_ISTRAP CNO Avails
- W TAB B3 - ISTRAP CNO Avails Fast Attack Subs_20241213 CLEAN
- W TAB C1 - DDG-FFG PY Acquisition Plan v9
- W TAB C2 - ASN(RDA) CR_TC_SEA_DD91-FFG PY ISTRAP
- W TAB C3 - ISTRAP_Integrated PY DDG 51-FFG 7_(DASN(P) signed_15Dec23
- W TAB D1 - OMNIBUS XI ISTRAP 392-23-120_For DASN(P) Signature
- W TAB D2 - ISTRAP Track Changes TAB A - OMNIBUS XI ISTRAP 392-23-120_Track Changes
- W TAB D3 - OMNIBUS XI Attack Submarine ISTRAP - Final 20231101

Initially the documents are reviewed against the NMCARS ANNEXES to ensure the submitter has followed the instructions and rules provided there

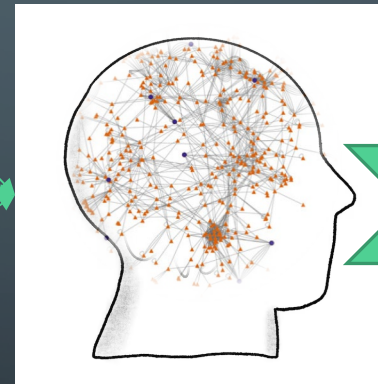
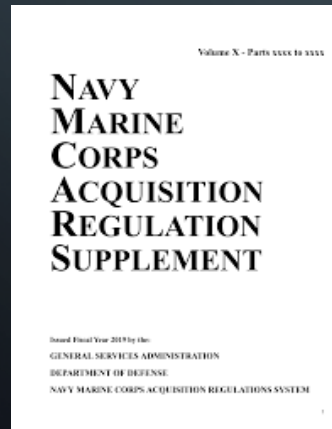


OUR SPECIFIC USE CASE, CONT'D

- TAB A - NMCARS 18-25 Annex 1 - J&A
- TAB B1 - CJA No. CR-24219 - Body - MS Word
- TAB B2 - CJA No. CR-24219 - Track Changes - DASN(P) and AGC
- TAB B3 - CJA No. CR-24219_Final Clean 22JAN25
- TAB C1 - Draper CPS JA23-51 Final CLEAN
- TAB C2 - 20240806_ASN(RD&A) CR_TC_Draper CPS JA23-51 Final CLEAN
- TAB C3 - Draper CPS JA23-51 Final CLEAN 9.10.2024
- TAB D1 - NAVSEA JA_DDG91_DMP2_to_NASSCO
- TAB D2 - ASN(RDA) CR_TC_NAVSEA_DDG91 DMP J&A - NASSCO - includes Legal Review
- TAB D3 - Final USS PINCKNEY (DDG 91) DMP JA 42.916_20231215

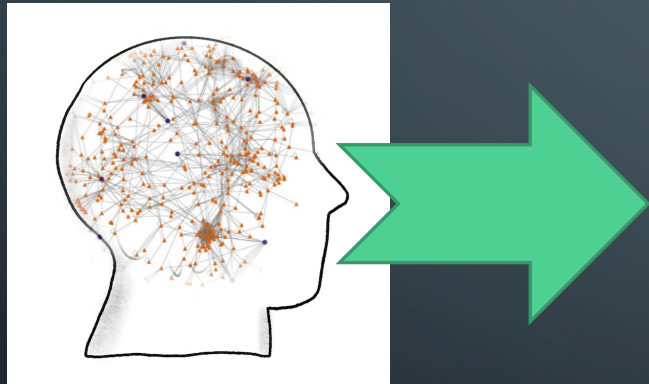
This structural review is somewhat rudimentary and repetitive for the staff to perform, they came to our team to see if it was something we could automate using a large language model

- TAB A - NMCARS 18-25 Annex 18 - ISTRAP
- TAB B1 - ISTRAP CNO Avails Fast Attack Subs
- TAB B2 - ASN(RDA) CR_TC_ISTRAP CNO Avails
- TAB B3 - ISTRAP CNO Avails Fast Attack Subs_20241213 CLEAN
- TAB C1 - DDG-FFG PY Acquisition Plan v9
- TAB C2 - ASN(RDA) CR_TC_SEA_DDG-FFG PY ISTRAP
- TAB C3 - ISTRAP Integrated PY DDG 51-FFG 7_(DASN(P) signed_15Dec23
- TAB D1 - OMNIBUS XI ISTRAP 392-23-120_For DASN(P) Signature
- TAB D2 - ISTRAP Track Changes TAB A - OMNIBUS XI ISTRAP 392-23-120_Track Changes
- TAB D3 - OMNIBUS XI Attack Submarine ISTRAP - Final 20231101



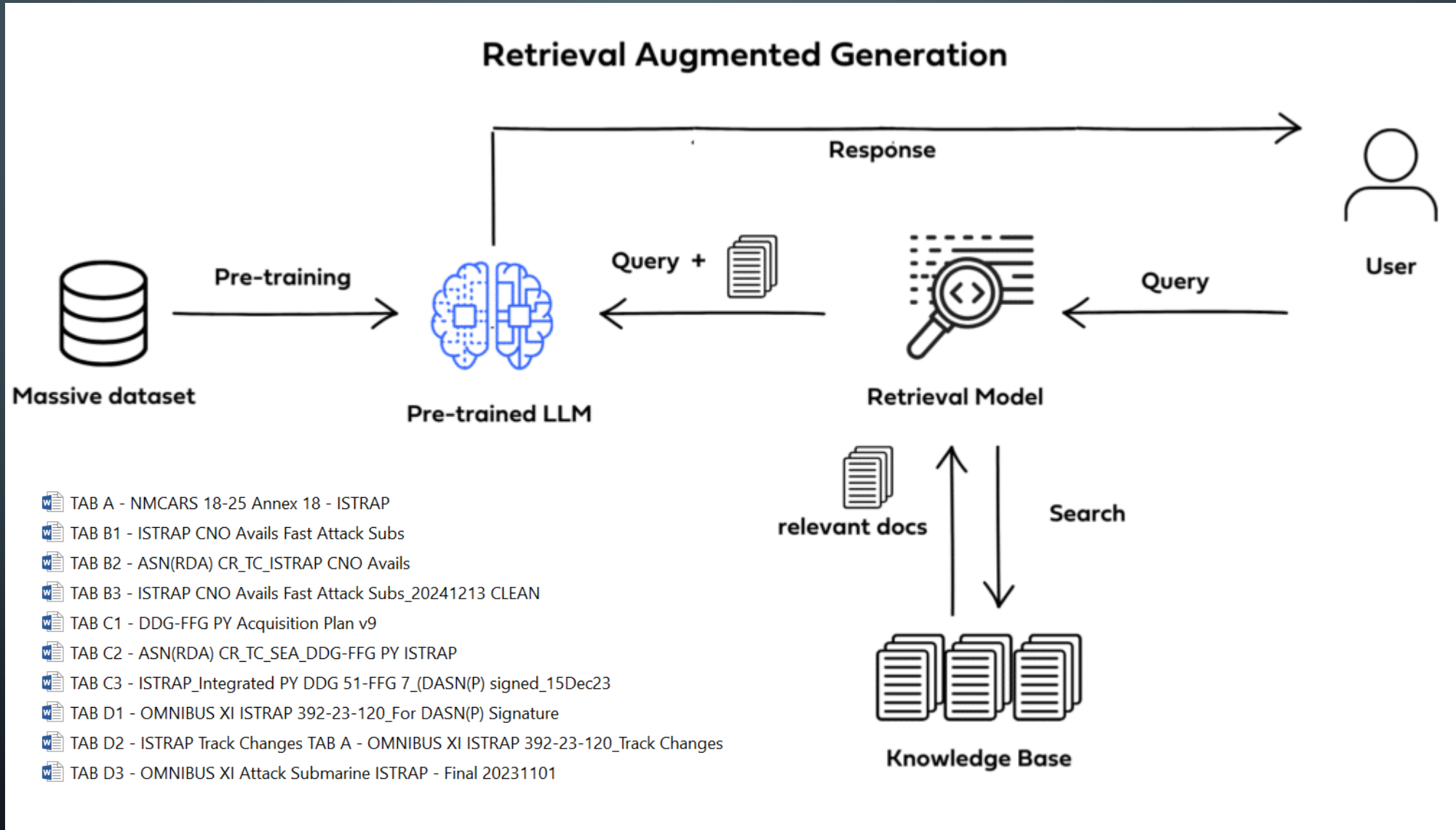
OUR SPECIFIC USE CASE, CONT'D

The results exceeded our expectations – we initially provided the model with first submissions from the field and asked it to assess the documents against NMCARS Annex Language which we also provided.



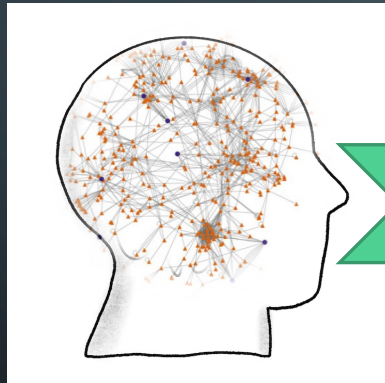
- The model noted the same discrepancies that our human reviewers found over 85% of the time
- The model found additional compliance issues that our human reviewers had not flagged
- The model presented specific recommendations to strengthen the documents from a logical argument standpoint

RAG - POWERING UP AI



RESULTS WITH RAG

By feeding the model our review team's specific feedback from past reviews, the model was able to better align to our expected outcomes, the match on comments increased by providing the specific historic record of our findings, however, the model could not 100% replicate the human review ...



Prompt 5: Tab C1

"Now that you know what the review team is looking for in a clean J&A, please review | the following document against the NMCARS Annex 1 requirements as though you were the review team:" I provided Tab C1

Response 5: Tab C1 Assessment

****Overall Assessment:**** This J&A is generally well-written and provides a comprehensive justification for sole-sourcing the contract to Draper. It addresses most of the key elements required by NMCARS Annex 1, but there are a few areas that could be further strengthened.

****Strengths:****

****Clear Description of the Requirement:**** Section 3 effectively outlines the supplies and services needed, including specific program requirements and a breakdown of

Assessment:** Both my review and the review team's comments J&A favorably, acknowledging its comprehensive justification for

* ****Focus on Section 5 (Rationale):**** Both reviews focused on the rationale provided in Section 5. We both agreed that it was generally strong but could be further improved.

* ****Market Research Acknowledgment:**** Both reviews acknowledged the thorough market research documented in Section 6.

****Differences:****

* ****Specific Questions vs. General Recommendations:**** My review provided more general recommendations for improvement (e.g., quantifying cost avoidance, clarifying | the "follow-on" aspect), whereas the review team asked specific questions related to contract length (base year plus four options), a seemingly extraneous asterisk, and the

Programmatic LLM Interface for quality

- Automates document review
 - Find areas for improvement
 - Highlight contracts that may not comply with NMCARS
- Model grades compliance with each NMCARS regulation



Overview of procedure

1. Provide LLM with contract and NMCARS regulations
2. Prompt LLM to output grades for point-by-point compliance
3. Summarize LLM outputs into human-readable interface

Chain-of-thought prompting

- Model gives reasoning before answers
- Allows model to “think through” the problem
- Improves answer quality

Standard Prompting

Input

Q: If a library had 30 shelves with 20 books on each shelf and someone added 10 more books on each shelf, how many books would there be in total?

Output

A: There are 600 books



Chain of Thought Prompting

Input

Q: If a library had 30 shelves with 20 books on each shelf, and someone added 10 more books on each shelf, how many books would there be in total?

Output

A: The library had 30 shelves with 20 books on each shelf, making a total of $30 \times 20 = 600$ books.

Ten books are being added to each shelf, for a total of $30 \times 10 = 300$ books.

So, it makes a total of $600 + 300 = 900$ books.



Hallucination Detection

- Addressing hallucinations with confidence scoring
- Certainty detection system based on model perplexity
- Low model confidence in answer (high perplexity) leads to flags for further review

Prediction error /
“surprise” at outcome

$$C(X) := 1 - PP(x) = 1 - \exp \sum_{i \in A:F} p_i \ln p_i$$

Probability
assigned to
grade i (A-F)

Example outputs from interface

Requirement str unique: 61	Grade str unique: 4	Certainty f64 
1.1: Statement of Need	A	42.16%
1.2: Historical Summary	A	46.88%
1.3.1: Program Objectives by Phase	A	72.34%
1.3.2: Contractor Data Requirements and Rights	B	46.65%
1.3.3: Cost Effectiveness of Buying Data	C	42.47%
1.3.4: Technical Data Package Validation	B	85.67%
1.3.5: Patents and Copyrights	B	76.33%
1.4: Funding Identification	A	14.3%
2.1.1: Product/Service Code Choice and Rationale	A	70.43%
2.1.2: Required Capabilities/Performance Standards	A	82.91%

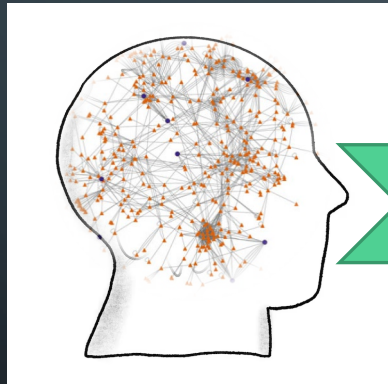
61 rows, 3 columns

Page 1 of 7 Download

Low certainty=
potential
hallucination!

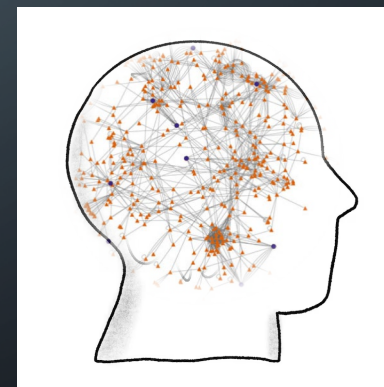
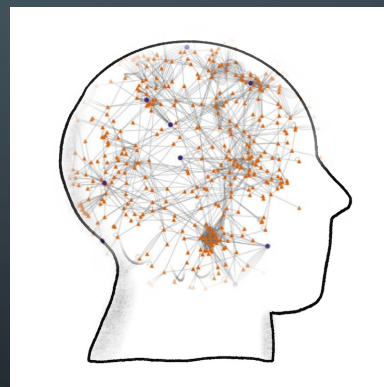
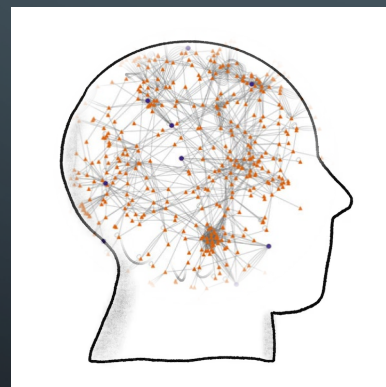
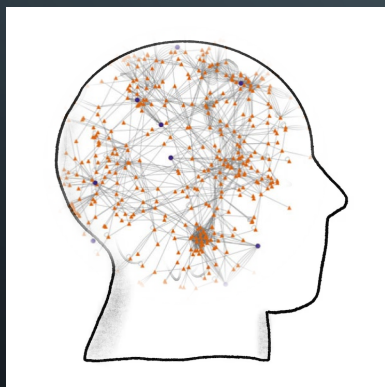
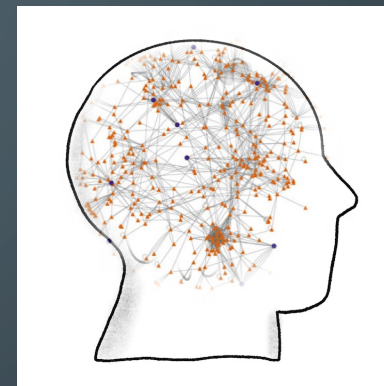
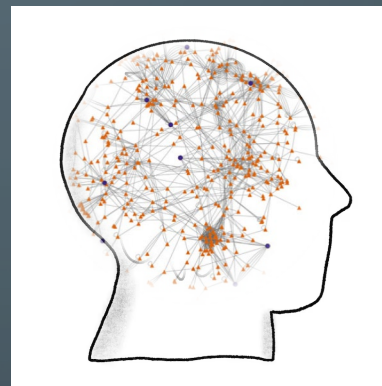
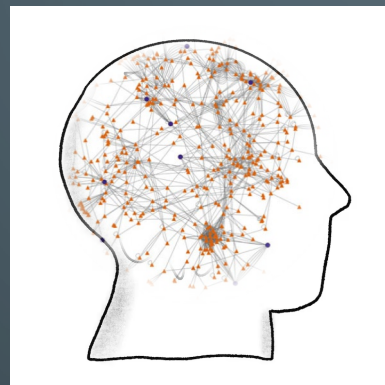
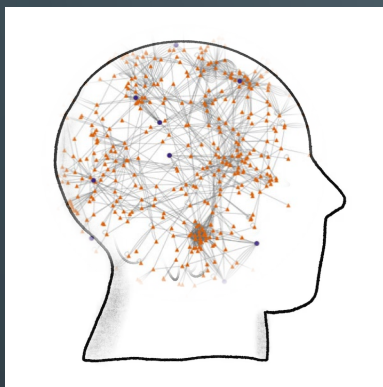
OVERALL FINDINGS

Large Language Models can 100% be used to supplement the reviews performed by the DASN(P) team with VERY positive results from our research



- Rudimentary structural reviews can be performed by our team or proactively by a supplementing pre-validation AI capability available at doc generation
- More complex logical analysis can also be performed by the AI, but definitely requires a “human in the loop” to combat model hallucination and increase certainty as well as adding professional guidance that the model cannot replicate

IMPACT ON ACQUISITION WORKFORCE



A BOLD NEW FUTURE...

